

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2006, 09:22:55 ; Search time 78.2895 Seconds
(without alignments)
39.286 Million cell updates/sec

Title: US-10-632-388-296
Perfect score: 37
Sequence: 1 ILAPPVP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	100.0	7	AAB17240	Aab17240 SH3 antag
2	37	100.0	7	ABB73233	Abb73233 Src homol
3	37	100.0	7	ADJ73387	Adj73387 SH3 antag
4	37	100.0	7	ADJ53021	Adj53021 CH1 delet
5	37	100.0	7	ADJ51982	Adj51982 CH1 delet
6	37	100.0	13	AAW05414	Aaw05414 Src SH3 d
7	37	100.0	13	AAW05481	Aaw05481 SH3-bindi
8	37	100.0	13	AAW11100	Aaw11100 Src SH3 d
9	37	100.0	13	AAW11117	Aaw11117 Src SH3 d
10	37	100.0	13	AAW38066	Aaw38066 PPPPY mot
11	37	100.0	13	AAW25513	Aaw25513 SH3 synth
12	37	100.0	13	ADB49256	Adb49256 Biotinyla
13	37	100.0	16	AAV24028	Aay24028 Synthetic
14	37	100.0	17	AAW05412	Aaw05412 Src SH3 d
15	37	100.0	30	AAW16938	Aaw16938 Random re
16	37	100.0	30	AAW25501	Aaw25501 Random pe
17	37	100.0	1767	ABB62726	Abb62726 Drosophil
18	36	97.3	16	AAW25387	Aaw25387 PLC gamma
19	36	97.3	137	ABP43530	Abp43530 Human sec
20	36	97.3	159	AAG83323	Aag83323 p patens
21	36	97.3	159	AAG80875	Aag80875 Lipid bio
22	36	97.3	200	ABB50640	Abb50640 Human sec
23	36	97.3	200	ABO44897	Abo44897 Novel hum
24	36	97.3	200	ABO26377	Abo26377 Protein a

98	33	89.2	1542	8	ADH71384	Adh71384 Human pro	171	32	86.5	1299	4	AAM78454	Aam78454 Human pro
99	33	89.2	1542	8	ADH71362	Adh71362 Human pro	172	32	86.5	1299	9	ADZ09816	Adz09816 Human bre
100	33	89.2	1542	8	ADH71370	Adh71370 Human pro	173	32	86.5	1307	4	AAM79438	Aam79438 Human pro
101	33	89.2	6310	6	ABU39869	Abu39869 Protein e	174	32	86.5	1307	4	AAM79439	Aam79439 Human pro
102	32	86.5	16	2	AAW25421	Aaw25421 Crk N-ter	175	32	86.5	1329	9	ADX07105	Adx07105 Cyclin-de
103	32	86.5	22	4	AAM16556	Aam16556 Peptide #	176	32	86.5	1630	4	ABB44550	Abb44550 Human wou
104	32	86.5	22	4	ABB35540	Abb35540 Peptide #	177	32	86.5	1630	6	ABP98331	Abp98331 Amino aci
105	32	86.5	22	4	AAM29038	Aam29038 Peptide #	178	32	86.5	1632	4	AAM39023	Aam39023 Human pol
106	32	86.5	22	4	ABB30366	Abb30366 Peptide #	179	32	86.5	2171	7	ADC99058	Adc99058 Human KPP
107	32	86.5	22	4	ABB20964	Abb20964 Protein #	180	32	86.5	2955	4	ABG221509	Abg221509 Novel hum
108	32	86.5	22	4	AAM68729	Aam68729 Human bon	181	32	86.5	2988	4	ABB44617	Abb44617 Human wou
109	32	86.5	22	4	AAM56353	Aam56353 Human bra	182	32	86.5	3530	8	ADQ18982	Adq18982 Human sof
110	32	86.5	22	4	ABG50398	Abg50398 Human liv	183	32	86.5	31267	6	ABG74786	Abg74786 Human RGS
111	32	86.5	22	4	AAM04269	Aam04269 Peptide #	184	32	86.5	34350	8	ADQ89964	Adq89964 Antagonis
112	32	86.5	22	5	ABG38314	Abg38314 Human pep	185	31	83.8	16	2	AAW25334	Aaw25334 Cortactin
113	32	86.5	28	3	AAy64765	Aay64765 Human 5'	186	31	83.8	20	2	AAW16957	Aaw16957 Cortactin
114	32	86.5	28	8	ADU72329	Adu72329 Signal pe	187	31	83.8	97	5	AAO21466	Aao21466 Eukaryoti
115	32	86.5	28	9	ADZ73320	Adz73320 Human inc	188	31	83.8	98	8	ADX73235	Adx73235 Plant ful
116	32	86.5	61	4	AAU60444	Aau60444 Propionib	189	31	83.8	108	4	ABG23885	Abg23885 Novel hum
117	32	86.5	61	6	ABM56963	Abm56963 Propionib	190	31	83.8	121	4	AAU46725	Aau46725 Propionib
118	32	86.5	80	8	ADX79019	Adx79019 Plant ful	191	31	83.8	121	6	ABM43244	Abm43244 Propionib
119	32	86.5	84	5	ABP00685	Abp00685 Human ORF	192	31	83.8	122	4	AAO09541	Aao09541 Human pol
120	32	86.5	101	5	ABP26410	Abp26410 Streptoco	193	31	83.8	132	5	ADK34787	Adk34787 Novel hum
121	32	86.5	101	8	ADV81495	Adv81495 Streptoco	194	31	83.8	143	7	ABO69656	AbO69656 Pseudomon
122	32	86.5	143	4	ABB71001	Abb71001 Drosophil	195	31	83.8	174	9	ABM94517	Abm94517 M. xanthu
123	32	86.5	151	7	ADE72523	Ade72523 Human end	196	31	83.8	177	7	ABO71055	AbO71055 Pseudomon
124	32	86.5	151	7	ADE72522	Ade72522 Human end	197	31	83.8	189	9	ABE39379	Aeb39379 L. pneumo
125	32	86.5	151	8	ADV88060	Adv88060 Streptoco	198	31	83.8	190	2	AAW88004	Aaw88004 Protein e
126	32	86.5	151	8	ADV79313	Adv79313 Streptoco	199	31	83.8	191	9	ABE35945	Aeb35945 L. pneumo
127	32	86.5	164	4	AAU42376	Aau42376 Propionib	200	31	83.8	269	8	ADS30919	Ads30919 Bacterial
128	32	86.5	164	6	ABM38895	Abm38895 Propionib	201	31	83.8	274	4	AAU42242	Aau42242 Propionib
129	32	86.5	217	9	ABM95613	Abm95613 M. xanthu	202	31	83.8	274	6	ABM38761	Abm38761 Propionib
130	32	86.5	282	3	AAG05722	Aag05722 Arabidops	203	31	83.8	302	8	ADY10393	Ady10393 Plant ful
131	32	86.5	297	3	AAG05721	Aag05721 Arabidops	204	31	83.8	310	2	AAY31953	Aay31953 Glutathio
132	32	86.5	315	7	ADC87269	Adc87269 Human GPC	205	31	83.8	310	7	ADK68446	Adk68446 Simian T-
133	32	86.5	331	3	AAG05720	Aag05720 Arabidops	206	31	83.8	310	8	ADU49881	Adu49881 StrLV-3 en
134	32	86.5	332	9	ADY30823	Ady30823 Thale cre	207	31	83.8	325	8	ADX68241	Adx68241 Plant ful
135	32	86.5	419	2	AAy50558	Aay50558 HSV2 LAT	208	31	83.8	332	3	AAG55461	Aag55461 Arabidops
136	32	86.5	419	7	ADG75126	Adg75126 Human her	209	31	83.8	341	3	AAG55460	Aag55460 Arabidops
137	32	86.5	433	5	AAW50962	Aaw50962 Maize met	210	31	83.8	347	2	AAW88005	Aaw88005 GA4 homol
138	32	86.5	433	5	AAW50942	Aaw50942 Maize met	211	31	83.8	347	5	ABB91678	Abb91678 Herbicida
139	32	86.5	433	6	ABU00021	Abu00021 Human nov	212	31	83.8	379	2	AAy31954	Aay31954 Glutathio
140	32	86.5	433	7	ADG19014	Adg19014 Maize ZmM	213	31	83.8	381	8	ADX78207	Adx78207 Plant ful
141	32	86.5	453	8	ADY06683	Ady06683 Plant ful	214	31	83.8	387	8	ADT58908	Adt58908 Plant pol
142	32	86.5	460	4	ABB71622	Abb71622 Drosophil	215	31	83.8	406	4	AAW95429	Aaw95429 Human pro
143	32	86.5	465	4	AAU46079	Aau46079 Propionib	216	31	83.8	406	4	AAU32015	Aau32015 Novel hum
144	32	86.5	465	6	ABM42598	Abm42598 Propionib	217	31	83.8	445	4	AAW94676	Aaw94676 Human pro
145	32	86.5	611	5	AAU78365	Aau78365 Arabidops	218	31	83.8	445	7	ADC31765	Adc31765 Human nov
146	32	86.5	611	9	ADY30825	Ady30825 Thale cre	219	31	83.8	445	8	ABM81200	Abm81200 Tumour-as
147	32	86.5	613	4	AAW94287	Aaw94287 Human pro	220	31	83.8	457	4	ABG02182	Abg02182 Novel hum
148	32	86.5	633	6	AAO16745	Aao16745 Aspergill	221	31	83.8	481	4	ABB63482	Abb63482 Drosophil
149	32	86.5	638	3	AAG38502	Aag38502 Arabidops	222	31	83.8	486	7	ADD47729	Add47729 Human Pro
150	32	86.5	659	7	ADC32976	Adc32976 Human nov	223	31	83.8	492	2	AAW85471	Aar85471 pTIV-L en
151	32	86.5	683	8	ADT55490	Adt55490 Plant pol	224	31	83.8	522	4	ABB70785	Abb70785 Drosophil
152	32	86.5	694	7	ADB64762	Adb64762 Human pro	225	31	83.8	524	6	ABU22998	Abu22998 Protein e
153	32	86.5	694	5	ABP63094	Abp63094 FLO11 gen	226	31	83.8	593	7	ABM89613	Abm89613 Rice abio
154	32	86.5	720	3	AAG38501	Aag38501 Arabidops	227	31	83.8	614	7	ABO71464	AbO71464 Pseudomon
155	32	86.5	740	6	ABJ25824	Abj25824 Aspergill	228	31	83.8	640	4	ABG16509	Abg16509 Novel hum
156	32	86.5	740	6	ABJ26424	Abj26424 Aspergill	229	31	83.8	854	9	ADX39652	Adx39652 HIV Env p
157	32	86.5	777	3	AAG38500	Aag38500 Arabidops	230	31	83.8	854	9	ADX39659	Adx39659 HIV Env p
158	32	86.5	779	8	ADX95214	Adx95214 Plant ful	231	31	83.8	856	9	ADX39650	Adx39650 HIV Env p
159	32	86.5	808	5	ABB05653	Abb05653 Human DNA	232	31	83.8	872	9	ADX39755	Adx39755 HIV Env p
160	32	86.5	885	8	ADP22490	Adp22490 Sea-squir	233	31	83.8	873	7	ADC07846	Adc07846 Rice prot
161	32	86.5	998	4	ABU53140	Abu53140 Human tes	234	31	83.8	949	4	ABB59399	Abb59399 Drosophil
162	32	86.5	1024	7	ADC31296	Adc31296 Human nov	235	31	83.8	1016	8	ADR08525	Adr08525 Human pro
163	32	86.5	1048	9	ADZ70351	Adz70351 Human pro	236	31	83.8	1027	7	ADJ69375	Adj69375 Human hea
164	32	86.5	1083	9	ADZ80792	Adz80792 Amino aci	237	31	83.8	1031	7	ABM88370	Abm88370 Rice abio
165	32	86.5	1100	4	AAW3286	Aaw3286 Human pro	238	31	83.8	1085	2	AAW90121	Aaw90121 Human huB
166	32	86.5	1100	6	AAO29563	Aao29563 Human Pci	239	31	83.8	1085	2	AAy41646	Aay41646 Human BUB
167	32	86.5	1100	9	ADY70491	Ady70491 Human bet	240	31	83.8	1085	3	AAy59142	Aay59142 Human ser
168	32	86.5	1192	8	ADN20849	Adn20849 Bacterial	241	31	83.8	1085	5	AAE25908	Aae25908 Human BUB
169	32	86.5	1258	4	AAW78455	Aaw78455 Human pro	242	31	83.8	1085	6	ABG75608	Abg75608 Human BUB
170	32	86.5	1299	3	AAy58633	Aay58633 Protein r	243	31	83.8	1085	7	ADG18109	Adg18109 Human kin

244	31	83.8	1085	7	ADG18095	Adg18095 Human kin	317	30	81.1	125	4	AAM68679	Aam68679 Human bon
245	31	83.8	1085	7	ADN95253	Adn95253 Human BEC	318	30	81.1	125	4	AAM56301	Aam56301 Human bra
246	31	83.8	1085	8	ADK67722	Adk67722 Human mod	319	30	81.1	125	4	ABG50343	Abg50343 Human liv
247	31	83.8	1085	8	ADO00999	Ado00999 Human hom	320	30	81.1	125	4	AAM04220	Aam04220 Peptide #
248	31	83.8	1085	8	ADO19728	Ado19728 Human PRO	321	30	81.1	125	5	ABG38258	Abg38258 Human pep
249	31	83.8	1085	8	ADQ09188	Adq09188 Human BUB	322	30	81.1	127	3	AAG26180	Aag26180 Zea mays
250	31	83.8	1085	9	ADY16259	Ady16259 PRO polyp	323	30	81.1	128	7	ADL06668	Adl06668 Human 3T3
251	31	83.8	1095	2	AAY06287	Aay06287 Human BUB	324	30	81.1	128	7	ADM05445	Adm05445 Human pro
252	31	83.8	1141	8	ADK60415	Adk60415 Angiogene	325	30	81.1	132	4	AAO07484	Aao07484 Human pol
253	31	83.8	1141	8	ADK60716	Adk60716 Angiogene	326	30	81.1	137	8	ADY08238	Ady08238 Plant ful
254	31	83.8	1141	8	ADP73339	Adp73339 Human mit	327	30	81.1	139	3	AAG23324	Aag23324 Arabidops
255	31	83.8	1175	4	ABG17468	Abg17468 Novel hum	328	30	81.1	139	3	AAG36753	Aag36753 Arabidops
256	31	83.8	1175	5	AAU76512	Aau76512 Human HPK	329	30	81.1	141	3	AAG23323	Aag23323 Arabidops
257	31	83.8	1175	8	ADK60186	Adk60186 Angiogene	330	30	81.1	142	4	AAU04823	Aau04823 Micromono
258	31	83.8	1175	8	ADK60487	Adk60487 Angiogene	331	30	81.1	144	4	AAM87652	Aam87652 Human imm
259	31	83.8	1175	8	ADP73110	Adp73110 Angiogene	332	30	81.1	145	4	AAU61889	Aau61889 Propionib
260	31	83.8	1212	4	AAE04368	Aae04368 Human kin	333	30	81.1	145	6	ABM58408	Abm58408 Propionib
261	31	83.8	1212	7	ADE25753	Ade25753 Human pro	334	30	81.1	147	3	AAG23322	Aag23322 Arabidops
262	31	83.8	1212	8	ADT98449	Adt98449 Human Mit	335	30	81.1	150	7	ABO79053	Abo79053 Pseudomon
263	31	83.8	1219	4	ABG17469	Abg17469 Novel hum	336	30	81.1	151	2	AAW54438	Aaw54438 Mouse nov
264	31	83.8	1230	2	AAW48895	Aaw48895 Candida a	337	30	81.1	151	2	AAY37644	Aay37644 Amino aci
265	31	83.8	1239	2	AAY55931	Aay55931 Human ZC1	338	30	81.1	151	3	AAB10245	Aab10245 Murine ad
266	31	83.8	1239	8	ADT98448	Adt98448 Human Mit	339	30	81.1	151	4	AAO03607	Aao03607 Human pol
267	31	83.8	1239	8	ADT98452	Adt98452 Human ger	340	30	81.1	154	3	AAG07339	Aag07339 Arabidops
268	31	83.8	1320	8	ADT98450	Adt98450 Human Mit	341	30	81.1	158	3	AAB41373	Aab41373 Human ORF
269	31	83.8	1384	8	ADN61482	Adn61482 Human KPP	342	30	81.1	160	8	ADR09710	Adr09710 Human pro
270	31	83.8	1392	8	ADJ96675	Adj96675 Human pro	343	30	81.1	163	4	AAM25236	Aam25236 Human pro
271	31	83.8	1655	4	ABG17466	Abg17466 Novel hum	344	30	81.1	163	4	ABB12270	Abb12270 Human gua
272	31	83.8	1665	4	ABB64010	Abb64010 Drosophil	345	30	81.1	163	5	ABU05941	Abu05941 M. tuberc
273	31	83.8	1726	7	ABM88576	Abm88576 Rice abio	346	30	81.1	163	6	ABM69378	Abm69378 Photorhab
274	31	83.8	2080	7	ADE31679	Ade31679 Human 707	347	30	81.1	164	7	ABO82761	Abo82761 Pseudomon
275	31	83.8	4881	3	AAB23751	Aab23751 S. avermi	348	30	81.1	167	4	ABG19252	Abg19252 Novel hum
276	31	83.8	4881	4	AAG65266	Aag65266 Streptomy	349	30	81.1	171	7	ABM89498	Abm89498 Rice abio
277	30	81.1	10	4	AAB81441	Aab81441 Target cy	350	30	81.1	177	8	ADP56387	Adp56387 Human PRO
278	30	81.1	24	7	ADB49304	Adb49304 Novel WW	351	30	81.1	178	8	ABO59894	Abo59894 Human gen
279	30	81.1	38	6	ABJ19800	Abj19800 DOCK 3 tu	352	30	81.1	198	4	AAB93283	Aab93283 Human pro
280	30	81.1	44	4	AAM88374	Aam88374 Human imm	353	30	81.1	200	8	ADX68196	Adx68196 Plant ful
281	30	81.1	52	4	ABG11384	Abg11384 Novel hum	354	30	81.1	201	8	ADY23460	Ady23460 Plant ful
282	30	81.1	62	4	AAU43820	Aau43820 Propionib	355	30	81.1	202	8	ABM82069	Abm82069 Tumour-as
283	30	81.1	62	6	ABM40339	Abm40339 Propionib	356	30	81.1	203	7	ADF04044	Adf04044 Bacterial
284	30	81.1	64	4	AAU45941	Aau45941 Propionib	357	30	81.1	204	3	AAG54583	Aag54583 Zea mays
285	30	81.1	64	6	ABM42460	Abm42460 Propionib	358	30	81.1	208	4	AAE01981	Aae01981 Human ATP
286	30	81.1	68	4	AAO02847	Aao02847 Human pol	359	30	81.1	214	8	ADX71182	Adx71182 Plant ful
287	30	81.1	69	4	AAU40618	Aau40618 Propionib	360	30	81.1	217	8	ADX96879	Adx96879 Plant ful
288	30	81.1	69	6	ABM37137	Abm37137 Propionib	361	30	81.1	218	2	AAW59198	Aaw59198 Seq ID 92
289	30	81.1	79	3	AAG02340	Aag02340 Human sec	362	30	81.1	218	2	AAW40087	Aaw40087 Seq ID 92
290	30	81.1	81	4	AAU47970	Aau47970 Propionib	363	30	81.1	219	2	AAW59197	Aaw59197 Seq ID 91
291	30	81.1	81	6	ABM44489	Abm44489 Propionib	364	30	81.1	219	2	AAW40086	Aaw40086 Seq ID 91
292	30	81.1	85	5	ABG62118	Abg62118 Human pro	365	30	81.1	226	7	ADI40526	Adi40526 Human pur
293	30	81.1	90	4	AAU39513	Aau39513 Propionib	366	30	81.1	226	8	ABO59343	Abo59343 Human gen
294	30	81.1	90	6	ABM36032	Abm36032 Propionib	367	30	81.1	231	7	ADC86879	Adc86879 Human GPC
295	30	81.1	93	5	ADK35932	Adk35932 Novel hum	368	30	81.1	232	4	ABG09954	Abg09954 Novel hum
296	30	81.1	98	4	AAO08196	Aao08196 Human pol	369	30	81.1	233	8	ADI42825	Adi42825 Plant tra
297	30	81.1	99	4	AAO00797	Aao00797 Human pol	370	30	81.1	233	8	ADO03021	Ado03021 Thalecres
298	30	81.1	100	2	AAY19684	Aay19684 SEQ ID NO	371	30	81.1	236	4	AAW42101	Aam42101 Human pol
299	30	81.1	101	7	ADM06137	Adm06137 Human pro	372	30	81.1	239	8	ADX88341	Adx88341 Plant ful
300	30	81.1	109	7	ADR42061	Adr42061 Human can	373	30	81.1	242	8	ADX72374	Adx72374 Plant ful
301	30	81.1	109	8	ADL83122	Adl83122 Human PRO	374	30	81.1	244	4	ABB62905	Abb62905 Drosophil
302	30	81.1	109	8	ADO19902	Ado19902 Human PRO	375	30	81.1	246	2	AAR29894	Aar29894 HCV NS4-N
303	30	81.1	109	8	ADP24582	Adp24582 PRO polyp	376	30	81.1	247	8	ADX76687	Adx76687 Plant ful
304	30	81.1	112	4	AAU45409	Aau45409 Propionib	377	30	81.1	250	6	AAO26516	Aao26516 Protein o
305	30	81.1	112	6	ABM41928	Abm41928 Propionib	378	30	81.1	255	3	AAB33114	Aab33114 Pinus rad
306	30	81.1	113	8	ABO58059	Abo58059 Human gen	379	30	81.1	257	3	AAB25424	Aab25424 Pinus rad
307	30	81.1	115	4	AAU50188	Aau50188 Propionib	380	30	81.1	260	7	ABM89772	Abm89772 Rice abio
308	30	81.1	115	5	ABP34174	Abp34174 Human ORF	381	30	81.1	263	5	ABP51376	Abp51376 Human MDD
309	30	81.1	115	6	ABM46707	Abm46707 Propionib	382	30	81.1	264	5	ABP43591	Abp43591 Human clo
310	30	81.1	122	3	AAG39025	Aag39025 Arabidops	383	30	81.1	264	8	ADT07510	Adt07510 Human col
311	30	81.1	122	8	ADX72660	Adx72660 Plant ,ful	384	30	81.1	268	7	ADG42188	Adg42188 Human bra
312	30	81.1	125	4	AAM16497	Aam16497 Peptide #	385	30	81.1	272	7	ABO80555	Abo80555 Pseudomon
313	30	81.1	125	4	ABB35483	Abb35483 Peptide #	386	30	81.1	273	4	ABG24064	Abg24064 Novel hum
314	30	81.1	125	4	AAM28985	Aam28985 Peptide #	387	30	81.1	273	7	ABO73472	Abo73472 Pseudomon
315	30	81.1	125	4	ABB30311	Abb30311 Peptide #	388	30	81.1	274	8	ADR67314	Adr67314 Human bla
316	30	81.1	125	4	ABB20920	Abb20920 Protein #	389	30	81.1	276	8	ADX92319	Adx92319 Plant ful

390	30	81.1	278	7	ADL07468	Adl07468	Wheat	TaD	463	30	81.1	481	7	ABO82916	Abo82916	Pseudomon
391	30	81.1	278	8	ADT59573	Adt59573	Plant	pol	464	30	81.1	486	8	ADX96724	Adx96724	Plant ful
392	30	81.1	278	8	ADX76499	Adx76499	Plant	ful	465	30	81.1	491	7	ABM85589	Abm85589	Human pro
393	30	81.1	279	7	ADY05913	Ady05913	Human	pro	466	30	81.1	495	7	ADY23100	Ady23100	Plant ful
394	30	81.1	279	8	ADY05913	Ady05913	Plant	ful	467	30	81.1	497	7	ADD14148	Add14148	Human src
395	30	81.1	281	4	ABB63034	Abb63034	Drosophil		468	30	81.1	524	7	ADJ70769	Adj70769	Human hea
396	30	81.1	282	4	AAU32765	Aau32765	Novel	hum	469	30	81.1	527	2	AAy49513	Aay49513	Xylulokin
397	30	81.1	285	8	ADX95838	Adx95838	Plant	ful	470	30	81.1	527	8	ADQ89930	Adq89930	Antagonis
398	30	81.1	286	8	ADX95837	Adx95837	Plant	ful	471	30	81.1	527	8	ADR66732	Adr66732	Human pro
399	30	81.1	287	9	AEA20252	Aea20252	Novel	hum	472	30	81.1	527	8	ADR66730	Adr66730	Human pro
400	30	81.1	288	8	ADY12581	Ady12581	Plant	ful	473	30	81.1	531	9	ADW08753	Adw08753	Human pro
401	30	81.1	288	8	ADY06457	Ady06457	Plant	ful	474	30	81.1	534	2	AAW25769	Aaw25769	Human pro
402	30	81.1	293	3	AAG06915	Aag06915	Arabidops		475	30	81.1	534	8	ADR58951	Adr58951	Human ELN
403	30	81.1	293	3	AAG28719	Aag28719	Arabidops		476	30	81.1	537	8	ADQ20678	Adq20678	Human sof
404	30	81.1	293	3	AAG40059	Aag40059	Arabidops		477	30	81.1	540	8	ADT57268	Adt57268	Plant pol
405	30	81.1	294	8	ADX78274	Adx78274	Plant	ful	478	30	81.1	541	9	AEA62712	Aea62712	Mitochond
406	30	81.1	294	8	ADX72801	Adx72801	Plant	ful	479	30	81.1	553	7	ABO68962	Abo68962	Pseudomon
407	30	81.1	296	3	AAG07338	Aag07338	Arabidops		480	30	81.1	560	7	ADC38726	Adc38726	Human sec
408	30	81.1	296	8	ADN74745	Adn74745	Thale cre		481	30	81.1	568	8	ADQ67693	Adq67693	Novel hum
409	30	81.1	297	9	AEA21066	Aea21066	Novel	hum	482	30	81.1	572	2	AAW31855	Aaw31855	Mycobacte
410	30	81.1	301	4	AAU36094	Aau36094	Klebsiell		483	30	81.1	575	8	ADT07559	Adt07559	Human col
411	30	81.1	301	6	ABU32009	Abu32009	Protein e		484	30	81.1	577	8	ADT07561	Adt07561	Human col
412	30	81.1	304	3	AAG06914	Aag06914	Arabidops		485	30	81.1	578	4	AAE01985	Aae01985	Human ATP
413	30	81.1	304	3	AAG28718	Aag28718	Arabidops		486	30	81.1	579	8	AAO24552	Aao24552	Murine BH
414	30	81.1	304	3	AAG40058	Aag40058	Arabidops		487	30	81.1	583	6	ABM66070	Abm66070	Propionib
415	30	81.1	304	4	ABB63066	Abb63066	Drosophil		488	30	81.1	583	9	ADW18471	Adw18471	Pinus rad
416	30	81.1	304	7	ABO66897	Abo66897	Klebsiell		489	30	81.1	585	7	ADT07557	Adt07557	Human col
417	30	81.1	304	8	ADY09972	Ady09972	Plant	ful	490	30	81.1	586	7	ADB65331	Adb65331	Human pro
418	30	81.1	305	2	AAR39356	Aar39356	Novel	hum	491	30	81.1	595	6	ABU33902	Abu33902	Protein e
419	30	81.1	305	8	ADN99683	Adn99683	Novel	hum	492	30	81.1	606	5	ABP43893	Abp43893	FLJ14692
420	30	81.1	312	8	ADO63087	Ado63087	Transcrip		493	30	81.1	609	5	AAU78364	Aau78364	Arabidops
421	30	81.1	312	8	ADN72439	Adn72439	Thale cre		494	30	81.1	609	9	ADY30827	Ady30827	Thale cre
422	30	81.1	312	9	AEA26203	Aea26203	Stress to		495	30	81.1	612	6	AAO27114	Aao27114	qSH-1 gen
423	30	81.1	317	4	ABG03344	Abg03344	Novel	hum	496	30	81.1	612	6	AAO27115	Aao27115	qSH-1 gen
424	30	81.1	318	4	AAAB92714	Aab92714	Human	pro	497	30	81.1	619	3	AAG41273	Aag41273	Arabidops
425	30	81.1	322	6	ADA13325	Ada13325	Human	int	498	30	81.1	623	4	AAE06678	Aae06678	Human nuc
426	30	81.1	330	8	ADS24352	Ads24352	Bacterial		499	30	81.1	623	7	ADC13533	Adc13533	Human nov
427	30	81.1	344	6	ADA06235	Ada06235	Wheat	cho	500	30	81.1	623	7	ADC31749	Adc31749	Human nov
428	30	81.1	344	8	ADO17006	Ado17006	Wheat	cho	501	30	81.1	625	8	ADS30749	Ads30749	Bacterial
429	30	81.1	346	3	AAG21361	Aag21361	Arabidops		502	30	81.1	629	6	ABU34518	Abu34518	Protein e
430	30	81.1	346	8	ADX72771	Adx72771	Plant	ful	503	30	81.1	634	8	ADQ18696	Adq18696	Human sof
431	30	81.1	353	7	ABM85588	Abm85588	Mouse	pro	504	30	81.1	634	8	ADY22469	Ady22469	Plant ful
432	30	81.1	353	8	ADY08875	Ady08875	Plant	ful	505	30	81.1	635	6	AAO26519	Aao26519	Protein o
433	30	81.1	354	8	ADN47354	Adn47354	Thermococ		506	30	81.1	638	8	ADX96261	Adx96261	Plant ful
434	30	81.1	364	3	AAG14403	Aag14403	Arabidops		507	30	81.1	640	8	ADS30049	Ads30049	Bacterial
435	30	81.1	365	3	AAG42876	Aag42876	Arabidops		508	30	81.1	641	3	AAG41272	Aag41272	Arabidops
436	30	81.1	366	8	ADY06751	Ady06751	Plant	ful	509	30	81.1	641	8	ADX92746	Adx92746	Plant ful
437	30	81.1	367	3	AAG14402	Aag14402	Arabidops		510	30	81.1	643	6	ABU36780	Abu36780	Protein e
438	30	81.1	367	6	ADA83987	Ada83987	Human	CDK	511	30	81.1	650	8	ADI82498	Adi82498	Human mod
439	30	81.1	368	3	AAG42875	Aag42875	Arabidops		512	30	81.1	650	8	ADY12848	Ady12848	Plant ful
440	30	81.1	375	7	ADE09101	Ade09101	Novel	pro	513	30	81.1	674	4	AAAB95253	Aab95253	Human pro
441	30	81.1	375	9	ADU40616	Adu40616	Novel	hum	514	30	81.1	676	7	ADC33302	Adc33302	Human nov
442	30	81.1	384	8	ADU02313	Adu02313	Novel	hum	515	30	81.1	703	8	ADH13194	Adh13194	Human mal
443	30	81.1	386	9	ADZ47759	Adz47759	Mycobacte		516	30	81.1	703	9	ADX07198	Adx07198	Cyclin-de
444	30	81.1	402	8	ADX91386	Adx91386	Plant	ful	517	30	81.1	703	9	AEA15081	Aea15081	Human pol
445	30	81.1	405	3	AAG21360	Aag21360	Arabidops		518	30	81.1	705	8	ADS28997	Ads28997	Bacterial
446	30	81.1	405	6	AAE36334	Aae36334	Arabidops		519	30	81.1	708	2	AAW36065	Aaw36065	Human neu
447	30	81.1	414	3	AAG21359	Aag21359	Arabidops		520	30	81.1	708	3	AAG41271	Aag41271	Arabidops
448	30	81.1	414	8	ADJ48846	Adj48846	Oil-assoc		521	30	81.1	708	7	ADE56151	Ade56151	Human pro
449	30	81.1	415	4	ABG30150	Abg30150	Novel	hum	522	30	81.1	708	7	ADE56159	Ade56159	Human pro
450	30	81.1	422	4	AAE01983	Aae01983	Human	ATP	523	30	81.1	708	7	ADE56155	Ade56155	Human pro
451	30	81.1	423	6	ABP99222	Abp99222	Orthosomy		524	30	81.1	708	8	ADJ64313	Adj64313	Cartilage
452	30	81.1	427	8	ADR86077	Adr86077	Aspergill		525	30	81.1	708	9	ADY14410	Ady14410	PRO polyp
453	30	81.1	428	4	ABG11496	Abg11496	Novel	hum	526	30	81.1	712	8	ADX92580	Adx92580	Plant ful
454	30	81.1	440	4	ABB61159	Abb61159	Drosophil		527	30	81.1	718	4	AAU41732	Aau41732	Propionib
455	30	81.1	442	4	ABB69463	Abb69463	Drosophil		528	30	81.1	718	4	ABG03655	Abg03655	Novel hum
456	30	81.1	445	5	ABP69609	Abp69609	Human	pol	529	30	81.1	718	6	ABM38251	Abm38251	Propionib
457	30	81.1	447	5	ABB05639	Abb05639	Hepatitis		530	30	81.1	735	4	ABB66233	Abb66233	Prosephil
458	30	81.1	449	2	AAW37129	Aaw37129	Hepatitis		531	30	81.1	739	6	ABU25591	Abu25591	Protein e
459	30	81.1	458	8	ADY08647	Ady08647	Plant	ful	532	30	81.1	743	4	ABB59515	Abb59515	Drosophil
460	30	81.1	476	3	AAB08405	Aab08405	Amino	aci	533	30	81.1	753	5	AAU82710	Aau82710	Amino aci
461	30	81.1	476	4	ABB62927	Abb62927	Drosophil		534	30	81.1	761	5	ABP74116	Abp74116	Human TRI
462	30	81.1	480	6	ABG74402	Abg74402	Green	pep	535	30	81.1	761	7	ADF74131	Adf74131	Human nov

536	30	81.1	763	2	AAW31852	Aaw31852 Mycobacte	609	30	81.1	1709	7	ADE06701	Ade06701 Hepaticis
537	30	81.1	763	6	ABU48371	Abu48371 Protein e	610	30	81.1	1788	4	AAU00016	Aau00016 Human Ple
538	30	81.1	779	8	ADT56006	Adt56006 Plant pol	611	30	81.1	1965	6	ABJ19786	Abj19786 DOCK 3 tu
539	30	81.1	810	9	ADZ22829	Adz22829 Muscle ca	612	30	81.1	1966	6	ABJ19797	Abj19797 DOCK 3 tu
540	30	81.1	814	3	AAB24089	Aab24089 Human PRO	613	30	81.1	1966	6	ABJ19789	Abj19789 DOCK 3 tu
541	30	81.1	814	4	ABG30224	Abg30224 Novel hum	614	30	81.1	1966	6	ABJ19794	Abj19794 DOCK 3 tu
542	30	81.1	814	5	AAE26669	Aae26669 Human cad	615	30	81.1	1966	6	ABJ19795	Abj19795 DOCK 3 tu
543	30	81.1	814	7	ADD14141	Add14141 Human src	616	30	81.1	1966	6	ABJ19796	Abj19796 DOCK 3 tu
544	30	81.1	814	8	ADP26909	Adp26909 Human M-c	617	30	81.1	1966	6	ABJ19785	Abj19785 DOCK 3 tu
545	30	81.1	814	8	ABM81804	Abm81804 Tumour-as	618	30	81.1	1966	6	ABJ19787	Abj19787 DOCK 3 tu
546	30	81.1	817	6	ABR53281	Abr53281 Protein s	619	30	81.1	1966	6	ABJ19798	Abj19798 DOCK 3 tu
547	30	81.1	817	7	ADK63588	Adk63588 Disease t	620	30	81.1	1966	6	ABJ19791	Abj19791 DOCK 3 tu
548	30	81.1	817	8	ADN19160	Adn19160 Bacterial	621	30	81.1	1966	6	ABJ19792	Abj19792 DOCK 3 tu
549	30	81.1	819	6	ABJ19348	Abj19348 NO VX rela	622	30	81.1	1966	6	ABJ19793	Abj19793 DOCK 3 tu
550	30	81.1	819	8	ADO41696	Ado41696 Novel hum	623	30	81.1	1966	6	ABJ19788	Abj19788 DOCK 3 tu
551	30	81.1	823	7	ABO68711	Abc68711 Pseudomon	624	30	81.1	1966	6	ABJ19790	Abj19790 DOCK 3 tu
552	30	81.1	875	4	ABB71072	Abb71072 Drosophil	625	30	81.1	1981	8	ADS11120	Ads11120 Human the
553	30	81.1	875	8	ADI40972	Adi40972 Fugu pher	626	30	81.1	1985	6	ABU09574	Abu09574 HCV Met-N
554	30	81.1	875	8	ADI41023	Adi41023 Fugu pher	627	30	81.1	1985	6	ABU09575	Abu09575 HCV Met-N
555	30	81.1	881	8	ADS30366	Ads30366 Bacterial	628	30	81.1	1985	6	ADR38450	Adr38450 Hepatitis
556	30	81.1	942	8	ADQ66974	Adq66974 Novel hum	629	30	81.1	2065	4	ABB63705	Abb63705 Drosophil
557	30	81.1	953	9	ABM95482	Abm95482 M. xanthu	630	30	81.1	2091	8	ADN61449	Adn61449 Human KPP
558	30	81.1	972	6	ABU25429	Abu25429 Protein e	631	30	81.1	2132	9	ADV97875	Adv97875 Murine pr
559	30	81.1	995	4	ABB62510	Abb62510 Drosophil	632	30	81.1	2193	6	ABR42219	Abr42219 Human pro
560	30	81.1	1052	4	ABG27366	Abg27366 Novel hum	633	30	81.1	2201	2	AAW01680	Aaw01680 HCV NS2-N
561	30	81.1	1056	4	AAE01980	Aae01980 Human ATP	634	30	81.1	2217	9	ADX98203	Adx98203 Lysine de
562	30	81.1	1082	7	ADE14368	Ade14368 Human int	635	30	81.1	2219	8	ADN00364	Adn00364 Novel hum
563	30	81.1	1085	8	ADS43269	Ads43269 Bacterial	636	30	81.1	2245	8	ADJ96651	Adj96651 Human Nim
564	30	81.1	1088	8	ADH61284	Adh61284 INTSIG pr	637	30	81.1	2382	6	ABP71619	Abp71619 Human WNK
565	30	81.1	1112	8	ADR66113	Adr66113 Human pro	638	30	81.1	2382	7	ADJ69598	Adj69598 Human hea
566	30	81.1	1112	8	ADR66455	Adr66455 Human pro	639	30	81.1	2382	9	ADX07450	Adx07450 Cyclin-de
567	30	81.1	1112	9	ADY18523	Ady18523 PRO polyp	640	30	81.1	2382	9	ADX98202	Adx98202 Lysine de
568	30	81.1	1116	6	ABJ26002	Abj26002 Aspergill	641	30	81.1	2388	8	ADS11119	Ads11119 Human the
569	30	81.1	1116	6	ABJ25402	Abj25402 Aspergill	642	30	81.1	3010	2	AAR20111	Aar20111 Non-A, no
570	30	81.1	1146	3	AA Y92225	Aay92225 Human pat	643	30	81.1	3010	2	AAR20091	Aar20091 Non-A, no
571	30	81.1	1203	2	AAY43261	Aay43261 Human pat	644	30	81.1	3010	2	AAR34580	Aar34580 Human hep
572	30	81.1	1203	2	AA Y28444	Aay28444 Human ptc	645	30	81.1	3010	2	AA Y06423	Aay06423 Non-A, no
573	30	81.1	1203	3	AA Y92703	Aay92703 Human pat	646	30	81.1	3010	9	ADX40788	Adx40788 HCV polym
574	30	81.1	1203	5	AAE19829	Aae19829 Human pat	647	30	81.1	3010	9	ADX40806	Adx40806 HCV polym
575	30	81.1	1203	6	ABG74104	Abg74104 Human pat	648	30	81.1	3011	2	AAR34468	Aar34468 Encoded b
576	30	81.1	1203	8	ADT07558	Adt07558 Human col	649	29	78.4	7	2	AAR84747	Aar84747 GST-dynam
577	30	81.1	1205	8	ADT07560	Adt07560 Human col	650	29	78.4	7	2	AAW79782	Aaw79782 Prolin-e-r
578	30	81.1	1213	8	ADT07556	Adt07556 Human col	651	29	78.4	7	3	AA Y69980	Aay69980 Src SH3 r
579	30	81.1	1234	4	ABB58721	Abb58721 Drosophil	652	29	78.4	8	2	AAR84745	Aar84745 DYN domai
580	30	81.1	1234	4	ABB58714	Abb58714 Drosophil	653	29	78.4	9	2	AAR84746	Aar84746 Dynamini r
581	30	81.1	1237	2	AAR13791	Aar13791 E75A prot	654	29	78.4	9	2	AAW99358	Aaw99358 Glycosyla
582	30	81.1	1237	6	AAE30117	Aae30117 Fruit fly	655	29	78.4	9	8	ADK88377	Adk88377 Human 191
583	30	81.1	1270	4	AAE01982	Aae01982 Human ATP	656	29	78.4	9	8	ADK88004	Adk88004 Human 191
584	30	81.1	1278	6	ABG74682	Abg74682 Human CGD	657	29	78.4	9	8	ADK88214	Adk88214 Human 191
585	30	81.1	1278	8	ADI82558	Adi82558 Human mod	658	29	78.4	9	8	ADK87474	Adk87474 Human 191
586	30	81.1	1345	5	AAE25097	Aae25097 Human kin	659	29	78.4	9	8	ADK87252	Adk87252 Human 191
587	30	81.1	1360	4	ABG30225	Abg30225 Novel hum	660	29	78.4	9	8	ADK83836	Adk83836 Human 191
588	30	81.1	1363	4	ABB64266	Abb64266 Drosophil	661	29	78.4	9	8	ADK87598	Adk87598 Human 191
589	30	81.1	1394	9	AEA32849	Aea32849 Modified	662	29	78.4	9	8	ADK88201	Adk88201 Human 191
590	30	81.1	1401	8	ADL13306	Adl13306 Human ste	663	29	78.4	9	8	ADK83353	Adk83353 Human 191
591	30	81.1	1401	9	ADX07508	Adx07508 Cyclin-de	664	29	78.4	9	8	ADK87960	Adk87960 Human 191
592	30	81.1	1419	8	ADN03630	Adn03630 Antipsori	665	29	78.4	9	8	ADK85972	Adk85972 Human 191
593	30	81.1	1426	4	AAE01984	Aae01984 Human ATP	666	29	78.4	9	8	ADK84372	Adk84372 Human 191
594	30	81.1	1426	5	ABP52159	Abp52159 Human 671	667	29	78.4	9	8	ADK86542	Adk86542 Human 191
595	30	81.1	1426	6	ABR54236	Abt54236 Human NOV	668	29	78.4	9	8	ADK84941	Adk84941 Human 191
596	30	81.1	1426	7	ADD37490	Add37490 Human tra	669	29	78.4	9	8	ADK85432	Adk85432 Human 191
597	30	81.1	1426	8	ADI27969	Adi27969 Human 671	670	29	78.4	9	8	ADK87823	Adk87823 Human 191
598	30	81.1	1426	9	ADY16951	Ady16951 PRO polyp	671	29	78.4	9	8	ADK87316	Adk87316 Human 191
599	30	81.1	1426	9	ADY20464	Ady20464 PRO polyp	672	29	78.4	9	8	ADK87734	Adk87734 Human 191
600	30	81.1	1426	9	ADY79907	Ady79907 Amino aci	673	29	78.4	9	8	ADK85509	Adk85509 Human 191
601	30	81.1	1445	8	ADQ66588	Adq66588 Novel hum	674	29	78.4	9	8	ADK87084	Adk87084 Human 191
602	30	81.1	1447	7	ABR57181	Abt57181 Drosophil	675	29	78.4	9	9	ADZ57264	Adz57264 Cytotoxic
603	30	81.1	1480	7	ADK63568	Adk63568 Disease-t	676	29	78.4	10	2	AAR93542	Aar93542 Random 10
604	30	81.1	1480	8	ADS43687	Ads43687 Bacterial	677	29	78.4	10	2	AAW70181	Aaw70181 N-termina
605	30	81.1	1538	4	AAB66466	Aab66466 Protein e	678	29	78.4	10	4	AAE05637	Aae05637 N-termina
606	30	81.1	1627	9	ABM92081	Abm92081 M. xanthu	679	29	78.4	10	6	ABO43501	Abo43501 M. tuberc
607	30	81.1	1647	4	ABB58428	Abb58428 Drosophil	680	29	78.4	10	7	ADF12073	Adf12073 Mycobacte
608	30	81.1	1709	7	ADE06699	Ade06699 Hepatitis	681	29	78.4	10	7	ADM40793	Adm40793 Mycobacte

682	29	78.4	10	8	ADK84706	Adk84706 Human 191	755	29	78.4	23	4	AAM29839	Aam29839 Peptide #
683	29	78.4	10	8	ADK85685	Adk85685 Human 191	756	29	78.4	23	4	ABB31143	Abb31143 Peptide #
684	29	78.4	10	8	ADK85777	Adk85777 Human 191	757	29	78.4	23	4	AAM69500	Aam69500 Human bon
685	29	78.4	10	8	ADK86765	Adk86765 Human 191	758	29	78.4	23	4	AAM57109	Aam57109 Human bra
686	29	78.4	10	8	ADK83626	Adk83626 Human 191	759	29	78.4	28	2	AAR84757	Aar84757 Dynamin r
687	29	78.4	10	8	ADK89039	Adk89039 Human 191	760	29	78.4	29	2	AAR92880	Aar92880 Mycobacte
688	29	78.4	10	8	ADK85168	Adk85168 Human 191	761	29	78.4	29	2	AAW75582	Aaw75582 M. tuberc
689	29	78.4	10	8	ADK89606	Adk89606 Human 191	762	29	78.4	29	6	ABG74421	Abg74421 M. tuberc
690	29	78.4	10	8	ADK89050	Adk89050 Human 191	763	29	78.4	30	2	AAR85676	Aar85676 45 kD M.t
691	29	78.4	10	8	ADK89421	Adk89421 Human 191	764	29	78.4	30	2	AAW18194	Aaw18194 N-termina
692	29	78.4	10	8	ADK88753	Adk88753 Human 191	765	29	78.4	30	6	ABU56346	Abu56346 M. tuberc
693	29	78.4	10	8	ADK89234	Adk89234 Human 191	766	29	78.4	30	7	AAE39289	Aae39289 M. tuberc
694	29	78.4	10	8	ADK86210	Adk86210 Human 191	767	29	78.4	30	7	ADF45131	Adf45131 M. tuberc
695	29	78.4	10	8	ADK84193	Adk84193 Human 191	768	29	78.4	30	8	ADO36816	Ado36816 45kD majo
696	29	78.4	10	8	ADK84663	Adk84663 Human 191	769	29	78.4	30	8	ADR30590	Adr30590 Human IGF
697	29	78.4	10	8	ADK84110	Adk84110 Human 191	770	29	78.4	30	8	ADU64191	Adu64191 Mycobacte
698	29	78.4	10	8	ADK89040	Adk89040 Human 191	771	29	78.4	34	4	AAM30735	Aam30735 Peptide #
699	29	78.4	11	2	AAW25461	Aaw25461 SH3 domai	772	29	78.4	34	4	ABB32023	Abb32023 Peptide #
700	29	78.4	11	2	AAW25477	Aaw25477 SH3 domai	773	29	78.4	34	4	AAM70405	Aam70405 Human bon
701	29	78.4	12	2	AAW25470	Aaw25470 SH3 domai	774	29	78.4	34	4	AAM57975	Aam57975 Human bra
702	29	78.4	12	2	AAW70109	Aaw70109 Peptide p	775	29	78.4	34	4	AAM05854	Aam05854 Peptide #
703	29	78.4	12	4	AAE05589	Aae05589 N-termina	776	29	78.4	37	2	AAR12910	Aar12910 Fragile X
704	29	78.4	12	4	AAE05570	Aae05570 N-termina	777	29	78.4	39	5	ABB77846	Abb77846 Amino aci
705	29	78.4	12	6	ABO43428	Abo43428 M. tuberc	778	29	78.4	43	4	AAM20313	Aam20313 Peptide #
706	29	78.4	12	6	ABO43452	Abo43452 M. tuberc	779	29	78.4	43	4	ABB40781	Abb40781 Peptide #
707	29	78.4	13	2	AAR84648	Aar84648 Grb2-SOS	780	29	78.4	43	4	AAM34547	Aam34547 Peptide #
708	29	78.4	14	2	AAR76702	Aar76702 N-termina	781	29	78.4	43	4	ABB24982	Abb24982 Protein #
709	29	78.4	14	2	AAW16953	Aaw16953 Src SH3 d	782	29	78.4	43	4	AAM74433	Aam74433 Human bon
710	29	78.4	15	2	AAW38946	Aaw38946 Peptide r	783	29	78.4	43	4	AAM61641	Aam61641 Human bra
711	29	78.4	15	8	ADK90145	Adk90145 Human 191	784	29	78.4	43	4	ABG56227	Abg56227 Human liv
712	29	78.4	15	8	ADK89791	Adk89791 Human 191	785	29	78.4	43	5	ABG44325	Abg44325 Human pep
713	29	78.4	15	8	ADK90444	Adk90444 Human 191	786	29	78.4	44	5	ABP28990	Abp28990 Streptoco
714	29	78.4	15	8	ADK89709	Adk89709 Human 191	787	29	78.4	47	7	ADH34263	Adh34263 Sterlet C
715	29	78.4	15	8	ADK90233	Adk90233 Human 191	788	29	78.4	48	7	ABM89624	Abm89624 Rice abio
716	29	78.4	15	8	ADK90020	Adk90020 Human 191	789	29	78.4	50	5	ABP34249	Abp34249 Human ORF
717	29	78.4	15	8	ADK89961	Adk89961 Human 191	790	29	78.4	52	4	AAU41114	Aau41114 Propionib
718	29	78.4	15	8	ADK89716	Adk89716 Human 191	791	29	78.4	52	6	ABM37633	Abm37633 Propionib
719	29	78.4	15	8	ADK90490	Adk90490 Human 191	792	29	78.4	54	4	AAU42763	Aau42763 Propionib
720	29	78.4	15	8	ADK90019	Adk90019 Human 191	793	29	78.4	54	6	ABM39282	Abm39282 Propionib
721	29	78.4	15	8	ADK90232	Adk90232 Human 191	794	29	78.4	58	5	ABP06010	Abp06010 Human ORF
722	29	78.4	15	8	ADK90234	Adk90234 Human 191	795	29	78.4	58	9	ADY64642	Ady64642 S. manson
723	29	78.4	15	8	ADK90388	Adk90388 Human 191	796	29	78.4	60	4	AAM82646	Aam82646 Human imm
724	29	78.4	15	8	ADK90470	Adk90470 Human 191	797	29	78.4	60	4	ABG20790	Abg20790 Novel hum
725	29	78.4	15	8	ADK89770	Adk89770 Human 191	798	29	78.4	64	8	ABO57564	Abo57564 Human gen
726	29	78.4	15	8	ADK90411	Adk90411 Human 191	799	29	78.4	67	8	ADK48110	Adk48110 Streptoco
727	29	78.4	16	2	AAW25364	Aaw25364 Abl SH3 d	800	29	78.4	70	4	AAU54512	Aau54512 Propionib
728	29	78.4	16	6	AAE36353	Aae36353 Human BSD	801	29	78.4	70	6	ABM51031	Abm51031 Propionib
729	29	78.4	17	2	AAR30076	Aar30076 N-termina	802	29	78.4	71	2	AAW75162	Aaw75162 Human sec
730	29	78.4	17	2	AAW05472	Aaw05472 SH3-bindi	803	29	78.4	71	2	AAW75163	Aaw75163 Human sec
731	29	78.4	17	2	AAW38051	Aaw38051 PPPPY mot	804	29	78.4	71	3	AAAY76073	Aay76073 Human ski
732	29	78.4	17	2	AAW32331	Aaw32331 Mycobacte	805	29	78.4	71	4	AAB56012	Aab56012 Skin cell
733	29	78.4	17	2	AAW32400	Aaw32400 Mycobacte	806	29	78.4	71	4	AAG77018	Aag77018 Human col
734	29	78.4	17	2	AAW64347	Aaw64347 Mycobacte	807	29	78.4	71	5	ABB72212	Abb72212 Human pro
735	29	78.4	17	2	AAW81714	Aaw81714 M. tuberc	808	29	78.4	71	6	ABO02039	Abo02039 Novel hum
736	29	78.4	17	2	AAW81690	Aaw81690 M. tuberc	809	29	78.4	71	6	ABO02038	Abo02038 Novel hum
737	29	78.4	17	2	AAAY39001	Aay39001 M. tuberc	810	29	78.4	71	9	ADZ12440	Adz12440 Human sec
738	29	78.4	17	2	AAAY38952	Aay38952 M. tuberc	811	29	78.4	71	9	ADZ12439	Adz12439 Human sec
739	29	78.4	17	2	AAAY39144	Aay39144 M. tuberc	812	29	78.4	73	4	AAU20656	Aau20656 Human sec
740	29	78.4	17	2	AAAY39090	Aay39090 M. tuberc	813	29	78.4	73	4	AAW91760	Aaw91760 Human imm
741	29	78.4	17	7	ADB49306	Adb49306 Novel ww	814	29	78.4	75	4	ABG26289	Abg26289 Novel hum
742	29	78.4	19	2	AAR48258	Aar48258 Nucleolar	815	29	78.4	76	9	ABM92044	Abm92044 M. xanthu
743	29	78.4	19	2	AAW96356	Aaw96356 Nucleolar	816	29	78.4	77	3	AAAY56033	Aay56033 Polypproli
744	29	78.4	19	5	ABG92991	Abg92991 Nucleolar	817	29	78.4	77	4	AAM89933	Aam89933 Human imm
745	29	78.4	19	6	ABP56586	Abp56586 Nucleus s	818	29	78.4	79	5	ABP34432	Abp34432 Human ORF
746	29	78.4	19	8	ADU07176	Adu07176 Nucleolar	819	29	78.4	86	5	ABP32728	Abp32728 Human ORF
747	29	78.4	19	8	ADT61875	Adt61875 Human nuc	820	29	78.4	87	4	ABG28592	Abg28592 Novel hum
748	29	78.4	19	9	ADW81324	Adw81324 Intrabody	821	29	78.4	87	8	ADG22729	Adg22729 Cyanophag
749	29	78.4	19	9	ADW88617	Adw88617 Nucleolus	822	29	78.4	90	3	AAB56590	Aab56590 Human pro
750	29	78.4	19	9	ADY32253	Ady32253 Novel can	823	29	78.4	91	4	AAU21866	Aau21866 Human car
751	29	78.4	19	9	AEB17233	Aeb17233 Human nuc	824	29	78.4	91	7	ADE45834	Ade45834 Human car
752	29	78.4	19	9	AEA43017	Aea43017 Nucleolar	825	29	78.4	91	8	ADJ07252	Adj07252 Human car
753	29	78.4	20	2	AAW16984	Aaw16984 ABL SH3 d	826	29	78.4	92	8	ABO57213	Abo57213 Human gen
754	29	78.4	22	9	AEA50220	Aea50220 IPF1 frag	827	29	78.4	95	4	AAM23515	Aam23515 Human EST

828	29	78.4	95	4	AAM85416	Aam85416 Human imm
829	29	78.4	95	4	AAM95844	Aam95844 Human rep
830	29	78.4	95	4	ABB96375	Abb96375 Human tes
831	29	78.4	98	4	AAO10464	Aao10464 Human pol
832	29	78.4	105	4	ABB15328	Abb15328 Human ner
833	29	78.4	105	9	AEA43772	Aea43772 Perhydrol
834	29	78.4	107	9	ABM90928	Abm90928 M. xanthu
835	29	78.4	110	8	ADM87746	Adm87746 Human EST
836	29	78.4	112	4	AAU39507	Aau39507 Propionib
837	29	78.4	112	6	ABM36026	Abm36026 Propionib
838	29	78.4	116	4	AAM19466	Aam19466 Peptide #
839	29	78.4	116	4	ABB38933	Abb38933 Peptide #
840	29	78.4	116	4	AAM32416	Aam32416 Peptide #
841	29	78.4	116	4	ABB23920	Abb23920 Protein #
842	29	78.4	116	4	AAU72154	Aau72154 Human bon
843	29	78.4	116	4	AAU46429	Aau46429 Propionib
844	29	78.4	116	4	AAM59584	Aam59584 Human bra
845	29	78.4	116	4	ABG53840	Abg53840 Human liv
846	29	78.4	116	5	ABG41968	Abg41968 Human pep
847	29	78.4	116	6	ABM42948	Abm42948 Propionib
848	29	78.4	117	4	AAM84403	Aam84403 Human imm
849	29	78.4	117	4	AAO10964	Aao10964 Human pol
850	29	78.4	119	5	ADK35157	Adk35157 Novel hum
851	29	78.4	121	4	AAB67557	Aab67557 Protein e
852	29	78.4	121	4	ABG26404	Abg26404 Novel hum
853	29	78.4	121	7	ADC333180	Adc333180 Human nov
854	29	78.4	121	7	ADC33316	Adc33316 Human nov
855	29	78.4	123	3	AAG27165	Aag27165 Zea mays
856	29	78.4	124	9	ABM97215	Abm97215 M. xanthu
857	29	78.4	125	2	AAV10797	Aay10797 Amino aci
858	29	78.4	125	7	ADB47915	Adb47915 Novel hum
859	29	78.4	125	8	ADJ55470	Adj55470 Novel hum
860	29	78.4	125	8	ADX87845	Adx87845 Plant ful
861	29	78.4	126	9	AEC04629	Aec04629 Human bre
862	29	78.4	127	4	AAO01796	Aao01796 Human pol
863	29	78.4	127	4	AAU31220	Aau31220 Novel hum
864	29	78.4	130	5	ABB89427	Abb89427 Human pol
865	29	78.4	130	8	ADQ65935	Adq65935 Novel hum
866	29	78.4	132	3	AAB53421	Aab53421 Human col
867	29	78.4	133	4	AAU50531	Aau50531 Propionib
868	29	78.4	133	4	ABG09963	Abg09963 Novel hum
869	29	78.4	133	5	ABU05968	Abu05968 M. tuberc
870	29	78.4	133	5	ABB98983	Abb98983 Endotheli
871	29	78.4	133	6	ABM47050	Abm47050 Propionib
872	29	78.4	133	6	ABU37006	Abu37006 Protein e
873	29	78.4	133	6	ABU34926	Abu34926 Protein e
874	29	78.4	133	7	ABO68139	Abo68139 Pseudomon
875	29	78.4	133	8	ADW05132	Adw05132 Human EDF
876	29	78.4	134	3	AAG26903	Aag26903 Zea mays
877	29	78.4	134	7	ABO68213	Abo68213 Pseudomon
878	29	78.4	135	7	ADM06091	Adm06091 Human pro
879	29	78.4	136	2	AAV12478	Aay12478 Human 5'
880	29	78.4	136	4	AAU51859	Aau51859 Propionib
881	29	78.4	136	6	ABM48378	Abm48378 Propionib
882	29	78.4	136	7	ABO81602	Abo81602 Pseudomon
883	29	78.4	137	4	AAU50475	Aau50475 Propionib
884	29	78.4	137	6	ABM46994	Abm46994 Propionib
885	29	78.4	138	4	AAO08702	Aao08702 Human pol
886	29	78.4	138	6	ABM65442	Abm65442 Propionib
887	29	78.4	139	4	AAU64104	Aau64104 Propionib
888	29	78.4	139	6	ABM60623	Abm60623 Propionib
889	29	78.4	142	6	ADA33033	Ada33033 Acinetoba
890	29	78.4	143	2	AAV36077	Aay36077 Extended
891	29	78.4	143	3	AAV65936	Aay65936 MUC1 muta
892	29	78.4	143	3	AAG44635	Aag44635 Zea mays
893	29	78.4	143	8	ADP19385	Adp19385 Human sec
894	29	78.4	144	7	ABO82491	Abo82491 Pseudomon
895	29	78.4	146	8	ABO55205	Abo55205 Human gen
896	29	78.4	150	8	ADY12206	Ady12206 Plant ful
897	29	78.4	151	6	ADA57406	Ada57406 Human sec
898	29	78.4	151	6	ADA41281	Ada41281 Human sec
899	29	78.4	151	6	ABR48027	Abr48027 Human sec
900	29	78.4	151	7	ADC74470	Adc74470 Human sec

901	29	78.4	151	7	ADD38035	Add38035 Human sec
902	29	78.4	153	2	AAW88293	Aaw88293 Wheat gem
903	29	78.4	155	8	ADR10386	Adr10386 Human pro
904	29	78.4	157	4	AAU62522	Aau62522 Propionib
905	29	78.4	157	6	ABM59041	Abm59041 Propionib
906	29	78.4	159	6	ABU33934	Abu33934 Protein e
907	29	78.4	159	6	ABU34553	Abu34553 Protein e
908	29	78.4	159	7	ADL36004	Adl36004 Human NOV
909	29	78.4	161	6	ABR56784	Abr56784 Human sec
910	29	78.4	161	8	ADY05803	Ady05803 Plant ful
911	29	78.4	162	3	AAG12442	Aag12442 Zea mays
912	29	78.4	162	4	AAU07336	Aau07336 1-aminocy
913	29	78.4	162	6	ABU35947	Abu35947 Protein e
914	29	78.4	162	6	ABU36794	Abu36794 Protein e
915	29	78.4	163	3	AAV65930	Aay65930 MUC1 muta
916	29	78.4	166	6	ADA11893	Ada11893 Human nov
917	29	78.4	166	8	ADX67667	Adx67667 Plant ful
918	29	78.4	167	2	AAV36228	Aay36228 Human sec
919	29	78.4	167	5	ABP42995	Abp42995 Human ova
920	29	78.4	167	6	ADA11722	Ada11722 Human nov
921	29	78.4	167	9	ABM97062	Abm97062 M. xanthu
922	29	78.4	170	4	AAO13099	Aao13099 Human pol
923	29	78.4	172	3	AAV65932	Aay65932 MUC1 muta
924	29	78.4	172	4	AAB52473	Aab52473 Mycobacte
925	29	78.4	172	9	ADW05076	Adw05076 C glutami
926	29	78.4	172	9	ABE91615	Aeb91615 Microbial
927	29	78.4	173	3	AAV65933	Aay65933 MUC1 muta
928	29	78.4	173	8	ADX78231	Adx78231 Plant ful
929	29	78.4	174	2	AAW64360	Aaw64360 Mycobacte
930	29	78.4	174	2	AAW81727	Aaw81727 M. tuberc
931	29	78.4	174	2	AAV39014	Aay39014 M. tuberc
932	29	78.4	174	2	AAV39157	Aay39157 M. tuberc
933	29	78.4	174	7	ABM88081	Abm88081 Rice abio
934	29	78.4	175	8	ADN99298	Adn99298 Novel hum
935	29	78.4	181	4	ABB50157	Abb50157 Human tra
936	29	78.4	183	3	AAG08187	Aag08187 Arabidops
937	29	78.4	183	7	ABM73994	Abm73994 DNA clone
938	29	78.4	185	7	ABO71429	Abo71429 Pseudomon
939	29	78.4	185	7	ABO77768	Abo77768 Pseudomon
940	29	78.4	186	8	ADX80177	Adx80177 Plant ful
941	29	78.4	187	2	AAW89878	Aaw89878 Antigen 1
942	29	78.4	187	3	AAG35655	Aag35655 Arabidops
943	29	78.4	188	2	AAW95491	Aaw95491 M. tuberc
944	29	78.4	189	4	AAU16302	Aau16302 Human nov
945	29	78.4	189	4	AAU34636	Aau34636 E. coli c
946	29	78.4	189	6	AAE30073	Aae30073 Arabidops
947	29	78.4	189	6	ABU55371	Abu55371 Human nov
948	29	78.4	189	9	ABE78057	Aeb78057 E. coli U
949	29	78.4	190	4	ABG08841	Abg08841 Novel hum
950	29	78.4	190	7	ABO70602	Abo70602 Pseudomon
951	29	78.4	191	7	ABO66030	Abo66030 Klebsiell
952	29	78.4	191	8	ADP04156	Adp04156 Human col
953	29	78.4	193	5	ABJ10473	Abj10473 Breast ca
954	29	78.4	193	5	ABP52841	Abp52841 Human ach
955	29	78.4	193	5	AAU10337	Aau10337 Novel hum
956	29	78.4	193	6	AAE33613	Aae33613 Human CAS
957	29	78.4	193	7	ADJ68845	Adj68845 Human hea
958	29	78.4	193	7	ADN39655	Adn39655 Cancer/an
959	29	78.4	193	8	ADM69070	Adm69070 Human TAT
960	29	78.4	193	9	ADX83185	Adx83185 Human TEG
961	29	78.4	195	6	ABM70751	Abm70751 Photorhab
962	29	78.4	200	8	ABO59883	Abo59883 Human gen
963	29	78.4	200	8	ADX87735	Adx87735 Plant ful
964	29	78.4	203	7	ABO82744	Abo82744 Pseudomon
965	29	78.4	204	8	ADX97034	Adx97034 Plant ful
966	29	78.4	206	8	ADX97250	Adx97250 Plant ful
967	29	78.4	207	9	ABM91766	Abm91766 M. xanthu
968	29	78.4	209	4	ABB67804	Abb67804 Drosophil
969	29	78.4	209	7	ABO63554	Abob63554 Klebsiell
970	29	78.4	210	7	ABO63406	Abob63406 Klebsiell
971	29	78.4	211	7	ADT88972	Adt88972 Tobacco M
972	29	78.4	213	4	ABG14699	Abg14699 Novel hum
973	29	78.4	213	7	ADB64730	Adb64730 Human pro

974 29 78.4 216 6 ADA48628 Ada48628 Rice prot
975 29 78.4 216 7 ADJ11720 Adj11720 Rice prot
976 29 78.4 216 7 ADJ11366 Adj11366 Rice prot
977 29 78.4 216 7 ABO75625 Abo75625 Pseudomon
978 29 78.4 217 5 ABJ10479 Abj10479 Breast ca
979 29 78.4 217 6 AAE33619 Aae33619 Human CAS
980 29 78.4 217 7 ABM88500 Abm88500 Rice abio
981 29 78.4 219 3 AAG35654 Aag35654 Arabidops
982 29 78.4 220 7 ABO67993 Abo67993 Pseudomon
983 29 78.4 221 6 ADA48512 Ada48512 Rice prot
984 29 78.4 221 7 ABR63182 Abr63182 Wheat pep
985 29 78.4 222 8 ADX91584 Adx91584 Plant ful
986 29 78.4 228 7 ABM87943 Abm87943 Rice abio
987 29 78.4 228 9 ABM96850 Abm96850 M. xanthu
988 29 78.4 229 8 ADT49803 Adt49803 Murine FB
989 29 78.4 230 8 ADT59535 Adt59535 Plant pol
990 29 78.4 231 4 AAB83411 Aab83411 Human ion
991 29 78.4 231 8 ADX71072 Adx71072 Plant ful
992 29 78.4 232 4 ABG11714 Abg11714 Novel hum
993 29 78.4 233 4 ABG16403 Abg16403 Novel hum
994 29 78.4 233 6 ABU19614 Abul19614 Protein e
995 29 78.4 235 6 ABR41859 Abr41859 Maize gro
996 29 78.4 239 8 ADX88356 Adx88356 Plant ful
997 29 78.4 242 8 ADX66752 Adx66752 Plant ful
998 29 78.4 243 2 AAR58815 Aar58815 Human c-m
999 29 78.4 243 3 AAG35653 Aag35653 Arabidops
1000 29 78.4 243 8 ADN73199 Adn73199 Thale cre

ALIGNMENTS

RESULT 1
AAB17240
ID AAB17240 standard; peptide; 7 AA.
XX
AC AAB17240;
XX
DT 31-OCT-2000 (first entry)
XX
DE SH3 antagonist peptide sequence SEQ ID NO:296.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CRLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 39; Page 299; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an

CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-P1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P³, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention

XX Sequence 7 AA;

Query Match 100.0%; Score 37; DB 3; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
Db 1 ILAPPVP 7

RESULT 2

ABB73233
ID ABB73233 standard; peptide; 7 AA.
XX
AC ABB73233;
XX
DT 05-APR-2002 (first entry)
XX
DE Src homology3 (SH3) antagonist peptide SEQ ID NO:296.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritis; antidiabetic; ophthalmological;
KW anastenic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX

PS Claim 39; Page 55; 176pp; English.

XX The present invention describes a vehicle-peptide molecule (I) or its

CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,

CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,

CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and

CC neuroprotective activities. (I) can be used as a therapeutic or

CC prophylactic agent as well as for screening purposes. (I) is useful for

CC diagnosing diseases characterised by dysfunction of their associated

CC protein of interest, for identifying normal or abnormal proteins of

CC interest, as a part of diagnostic kit to detect the presence of their

CC proteins of interest in a biological sample. Additionally, (I) is useful

CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,

CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,

CC infertility, and neurological degenerative diseases. (I), comprising EPO-

CC mimetic compounds are useful for treating disorders characterised by low

CC red blood cell levels such as anaemia. The TPO-mimetic comprising

CC compounds are useful for treating conditions that involve an existing

CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet

CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic

CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,

CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777

CC represent amino acid and nucleic acid sequences used in the

CC exemplification of the present invention

XX

SQ Sequence 7 AA;

Query Match 100.0%; Score 37; DB 5; Length 7;

Best Local Similarity 100.0%; Pred. No. 2e+06;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7

Db |||||

1 ILAPPVP 7

RESULT 3

ADJ73387

ID ADJ73387 standard; peptide; 7 AA.

XX

AC ADJ73387;

XX

DT 06-MAY-2004 (first entry)

XX

DE SH3 antagonist peptide sequence SeqID 842.

XX

KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;

KW cardiovascular; infectious; malignant; neurologic disease; anaemia;

KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;

KW SH3.

XX

OS Synthetic.

XX

PN WO2003084477-A2.

XX

PD 16-OCT-2003.

XX

PF 24-MAR-2003; 2003WO-US009139.

XX

PR 29-MAR-2002; 2002US-0368791P.

XX

PA (CENZ) CENTOCOR INC.

XX

PI Heavner GA, Knight DM, Scallion BJ, Ghrayeb J;

XX

DR WPI; 2003-804237/75.

XX

PT New CDR mimetibody comprising a portion of a heavy or light chain

PT variable region comprising human framework or ligand binding region,

PT useful for preparing a composition for treating e.g., immune,

PT cardiovascular or neurologic disease.

XX

PS Disclosure; SEQ ID NO 842; 97pp; English.

XX This invention relates to novel mammalian CDR mimetibodies, specific

CC portions or variants thereof. Specifically, it refers to an antibody

CC fragment where a protein has been inserted into, or replaces a portion

CC of, one or more CDR regions, such that each CDR mimetibody comprises at

CC least one portion of a heavy chain or light chain variable region, which

CC itself comprises at least one human framework region and at least one

CC ligand binding region (LBR). The present invention describes human

CC mimetibodies, including modified immunoglobulins and cleavage products

CC that can be useful in gene therapy and the generation of transgenic

CC plants and animals. Furthermore, the CDR mimetibody is useful for

CC preparing compositions for modulating, treating or reducing the symptoms

CC of immune, cardiovascular, infectious, malignant and/ or neurologic

CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,

CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This

CC peptide sequence is an SH3 antagonist peptide sequence used to make a

CC mimetibody of the invention.

XX

SQ Sequence 7 AA;

Query Match 100.0%; Score 37; DB 7; Length 7;

Best Local Similarity 100.0%; Pred. No. 2e+06;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7

Db |||||

1 ILAPPVP 7

RESULT 4

ADJ53021

ID ADJ53021 standard; peptide; 7 AA.

XX

AC ADJ53021;

XX

DT 06-MAY-2004 (first entry)

XX

DE CH1 deleted mimetibody-related peptide SeqID842.

XX

KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;

KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;

KW fungicide; gene therapy; immune disorder; cardiovascular disease;

KW arrhythmia; hypertension; heart failure; neurodegenerative;

KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;

KW cancerous condition; infectious disease; bacterial infection;

KW viral infection; fungal infection.

XX

OS Unidentified.

OS Synthetic.

XX

PN WO2004002417-A2.

XX

PD 08-JAN-2004.

XX

PF 27-JUN-2003; 2003WO-US020347.

XX

PR 28-JUN-2002; 2002US-0392431P.

XX

PA (CENZ) CENTOCOR INC.

XX

PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;

PI Kutoloski KA,

XX

DR WPI; 2004-082870/08.

XX

PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for

PT modulating, treating, alleviating, preventing an immune, cardiovascular,

PT or neurodegenerative disease or disorder, anemia, cancer, or infectious

PT diseases.

XX

PS Claim 3; SEQ ID NO 842; 129pp; English.

XX

CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences

CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.

XX Sequence 7 AA;

Query Match 100.0%; Score 37; DB 8; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
Db 1 ILAPPVP 7

RESULT 5

ADJ51982
ID ADJ51982 standard; peptide; 7 AA.

XX ADJ51982;

XX 06-MAY-2004 (first entry)

DE CH1 deleted mimetibody-related peptide SeqID842.

XX CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmologic; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.

XX Unidentified.
OS Synthetic.

XX WO2004002424-A2.

PN 08-JAN-2004.

PF 30-JUN-2003; 2003WO-US020495.

PR 28-JUN-2002; 2002US-0392431P.

PR 19-SEP-2002; 2002US-0412144P.

XX (CENZ) CENTOCOR INC.

XX Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;

XX WPI; 2004-082872/08.

XX New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and

PT nutritional disorders.

XX Claim 15; SEQ ID NO 842; 123pp; English.

XX This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC antiinflammatory, neuroleptic, ophthalmologic, hepatotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CH1 deleted mimetibody, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetibody of the invention.

XX Sequence 7 AA;

Query Match 100.0%; Score 37; DB 8; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7

Db 1 ILAPPVP 7

RESULT 6

AAW05414
ID AAW05414 standard; peptide; 13 AA.

XX AAW05414;

XX 24-FEB-1998 (first entry)

DE Src SH3 domain-binding peptide, T12SRC.1.

XX Src-homology region 3 domain; human; mouse; SH3 domain; cell growth;
KW cellular signalling element; cellular structural element; malignancy;
KW protein identification; functional domain; protein screening;
KW cellular signal transduction process; binding peptide.

OS Homo sapiens.
OS Synthetic.

XX Key Location/Qualifiers

FH Modified-site 1 /note= "Biotin labelled"

FT Modified-site 13 /note= "C-terminal amide"

XX WO9631625-A1.

XX 10-OCT-1996.

XX 04-APR-1996; 96WO-US004454.

XX 07-APR-1995; 95US-00417872.

PR 03-APR-1996; 96US-00630915.

XX (CYTO-) CYTOGEN CORP.

PA (UYNC-) UNIV NORTH CAROLINA.

XX Sparks AB, Hoffman N, Kay BK, Fowlkes DM, McConnell SJ;

DR WPI; 1996-465045/46.

XX Identifying polypeptide(s) having specific functional domain (esp. SH3

PT domain) - comprises detecting selective binding to recognition unit,

PT regardless of sequence homology.

XX Example; Page 81; 174pp; English.

PS

XX AAW05414 and AAW05415 represent variants of the Src Src-homology region 3

CC (SH3) domain-binding peptide termed pSrcCII (see AAW05412). These

CC sequences were used to probe human cDNA libraries to identify human SH3

CC domain containing proteins (such as AAW05400), that can be used in the

CC method of the invention. The method of the invention is for identifying

CC polypeptides containing functional domains of interest (especially SH3

CC domains). It comprises contacting a multivalent recognition unit (RU)

CC complex with a number of peptides and identifying polypeptides having a

CC selective binding affinity for the RU complex. The method is based on

CC functional similarities and does not rely on sequence similarities. Prior

CC methods only gave limited success for identifying proteins containing an

CC SH3 domain due to the minimal sequence homology among known SH3 proteins.

CC Multivalent RU complexes are particularly suited to screening for

CC polypeptides containing functional domains that are similar to, but not

CC identical in sequence to, the original target functional domain. The new

CC method enables proteins having a common function to be identified.

CC Identification of novel SH3 proteins will be useful for a better

CC understanding of cell growth, malignancy, signal transduction processes,

CC etc. New candidate drugs can be identified, and their specificities (e.g.

CC pharmacological activities) can be assessed using the method of the

CC invention

XX

SQ Sequence 13 AA;

Query Match 100.0%; Score 37; DB 2; Length 13;

Best Local Similarity 100.0%; Pred. No. 28;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7

Db 2 ILAPPVP 8

RESULT 7

AAW05481

ID AAW05481 standard; peptide; 13 AA.

XX

AC AAW05481;

XX

DT 24-FEB-1998 (first entry)

XX

DE SH3-binding peptide T12SRC1.

XX

KW Src-homology region 3 domain; human; mouse; SH3 domain; cell growth;

KW cellular signalling element; cellular structural element; malignancy;

KW protein identification; functional domain; protein screening;

KW cellular signal transduction process; binding peptide.

XX

OS Synthetic.

XX

PN WO9631625-A1.

XX

PD 10-OCT-1996.

XX

PF 04-APR-1996; 96WO-US004454.

XX

PR 07-APR-1995; 95US-00417872.

PR 03-APR-1996; 96US-00630915.

XX

PA (CYTO-) CYTOGEN CORP.

PA (UYNC-) UNIV NORTH CAROLINA.

XX

PI Sparks AB, Hoffman N, Kay BK, Fowlkes DM, McConnell SJ;

XX WPI; 1996-465045/46..

DR

XX Identifying polypeptide(s) having specific functional domain (esp. SH3

PT domain) - comprises detecting selective binding to recognition unit,

PT regardless of sequence homology.

XX Example; Fig 13; 174pp; English.

PS

XX AAW05445-W05492 represent Src-homology region 3 (SH2) domain binding

CC peptides. These sequences were used as parts of multivalent recognition

CC unit complexes used in the method of the invention. The method of the

CC invention is for identifying polypeptides containing functional domains

CC of interest (especially SH3 domains). It comprises contacting a

CC multivalent recognition unit (RU) complex with a number of peptides and

CC identifying polypeptides having a selective binding affinity for the RU

CC complex. The method is based on functional similarities and does not rely

CC on sequence similarities. Prior methods only gave limited success for

CC identifying proteins containing an SH3 domain due to the minimal sequence

CC homology among known SH3 proteins. Multivalent RU complexes are

CC particularly suited to screening for polypeptides containing functional

CC domains that are similar to, but not identical in sequence to, the

CC original target functional domain. The new method enables proteins having

CC a common function to be identified. Identification of novel SH3 proteins

CC will be useful for a better understanding of cell growth, malignancy,

CC signal transduction processes, etc. New candidate drugs can be

CC identified, and their specificities (e.g. pharmacological activities) can

CC be assessed using the method of the invention

XX

SQ Sequence 13 AA;

Query Match 100.0%; Score 37; DB 2; Length 13;

Best Local Similarity 100.0%; Pred. No. 28;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7

Db 2 ILAPPVP 8

RESULT 8

AAW11100

ID AAW11100 standard; peptide; 13 AA.

XX

AC AAW11100;

XX

DT 25-JUN-1997 (first entry)

XX

DE Src SH3 domain-binding peptide used in signal transduction modulation.

XX

KW Src; SH3; Src homology region 3; binding affinity; oncogenic protein;

KW protein tyrosine kinase; signal transduction; RNA processing;

KW trafficking; translation.

XX

OS Synthetic.

XX

PN WO9603649-A1.

XX

PD 08-FEB-1996.

XX

PF 24-JUL-1995; 95WO-US009382.

XX

PR 22-JUL-1994; 94US-00278865.

PR 07-JUN-1995; 95US-00483555.

XX

PA (UYNC-) UNIV NORTH CAROLINA.

XX

PI Sparks AB, Kay BK, Thorn JM, Quilliam LA, Der CJ;

XX WPI; 1996-117151/12.

DR

XX Peptide with binding affinity for Src homology region 3 (SH3) domains of

PT proteins - useful for e.g. modulating signal transduction pathways at the

PT cellular level, esp. protein tyrosine kinase-mediated.

XX

PS Claim 34; Page 80; 116pp; English.
XX AAW11098-W11124 are peptides that bind to the Src SH3 domain. The SH3
CC binding peptides are useful in modulating signal transduction pathways at
CC the cellular level (especially protein tyrosine kinase-mediated),
CC modulating oncogenic protein activity, or providing compounds for the
CC development of drugs with the ability to modulate broad classes, as well
CC as specific classes, of proteins involved in signal transduction and also
CC for regulating the processing, trafficking or translation of RNA.
CC Conjugates of the peptides with detectable labels or imaging agents are
CC useful for imaging cells, tissues and organs in which Src or Src-related
CC proteins are expressed
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 37; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
| | | | |
Db 2 ILAPPVP 8

RESULT 9
AAW11117
ID AAW11117 standard; peptide; 13 AA.
XX
AC AAW11117;
XX
DT 25-JUN-1997 (first entry)
XX
DE Src SH3 domain-binding peptide used in signal transduction modulation.
XX
KW Src; SH3; Src homology region 3; binding affinity; oncogenic protein;
KW protein tyrosine kinase; signal transduction; RNA processing;
KW trafficking; translation.
XX
OS Synthetic.
XX
PN WO9603649-A1.
XX
PD 08-FEB-1996.
XX
PF 24-JUL-1995; 95WO-US009382.
XX
PR 22-JUL-1994; 94US-00278865.
PR 07-JUN-1995; 95US-00483555.
XX
PA (UYNC-) UNIV NORTH CAROLINA.
XX
PI Sparks AB, Kay BK, Thorn JM, Quilliam LA, Der CJ;
XX
DR WPI; 1996-117151/12.
XX
PT Peptide with binding affinity for Src homology region 3 (SH3) domains of
PT proteins - useful for e.g. modulating signal transduction pathways at the
PT cellular level, esp. protein tyrosine kinase-mediated.

PS Claim 39; Page 85; 116pp; English.
XX
CC AAW11098-W11124 are peptides that bind to the Src SH3 domain. The SH3
CC binding peptides are useful in modulating signal transduction pathways at
CC the cellular level (especially protein tyrosine kinase-mediated),
CC modulating oncogenic protein activity, or providing compounds for the
CC development of drugs with the ability to modulate broad classes, as well
CC as specific classes, of proteins involved in signal transduction and also
CC for regulating the processing, trafficking or translation of RNA.
CC Conjugates of the peptides with detectable labels or imaging agents are
CC useful for imaging cells, tissues and organs in which Src or Src-related
CC proteins are expressed
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 37; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ILAPPVP 7
| | | | |
Db 4 ILAPPVP 10

RESULT 10
AAW38066
ID AAW38066 standard; peptide; 13 AA.
XX
AC AAW38066;

XX
DT 23-APR-1998 (first entry)

XX PPPPY motif containing peptide used to bind WW domains.

XX
KW Peptide recognition unit; WW domain; cell signalling; growth regulation;
KW cytoskeleton organisation; targeted drug screening; modulator;
KW WW domain interaction; YAP protein; dystrophin.

XX
OS Synthetic.

XX
PN WO9737223-A1.

XX
PD 09-OCT-1997.

XX
PF 03-APR-1997; 97WO-US005547.

XX
PR 03-APR-1996; 96US-00630916.

XX
PA (CYTO-) CYTOGEN CORP.

XX
PA (UYNC-) UNIV NORTH CAROLINA.

XX
PI Pirozzi G, Kay BK, Fowlkes DM;

XX
DR WPI; 1997-503234/46.

XX
PT Identifying cell signalling and growth regulatory polypeptides by
PT reaction with multivalent recognition complex - polypeptides are useful
PT in targeted drug selection.

XX
PS Disclosure; Fig 15A; 220pp; English.

XX
CC Peptides AAW38057-67 contain PPPPY-like motifs. The PPPY motif is found
CC in the proline rich regions of WBP-1 and WBP-2 proteins. Peptides
CC containing this residue have been shown to bind the YAP WW domain, but
CC not the WW domain from dystrophin or to a panel of SH3 domains. Peptides
CC AAW38057-67 were biotinylated and complexed with alkaline streptavidin,
CC and used in a cross affinity mapping experiment. They were tested for
CC their ability to bind to the 12 individual novel WW domains of WWP1
CC (AAW36794), WWP2 (AAW36795), WWP3 (AAW37696) and WWP4 (AAW36797), which
CC were expressed as glutathione-S-transferase expression proteins. The
CC present peptide, derived from Src protein, does not bind the WW domains
CC of the novel proteins. The WW domain is a small functional domain. Its
CC name is derived from the observation that two tryptophan residues, one in
CC the amino terminal portion of the WW domain and one in the carboxyl
CC terminal portion, are conserved. Most proteins containing WW domains have
CC a function involving cell signalling and growth regulation or the
CC organisation of the cytoskeleton. Polypeptides containing a WW domain are
CC identified by treating a multivalent recognition unit complex that has
CC selective binding affinity for a WW domain, with many polypeptides and
CC identifying those with selective affinity for the complex. Proteins
CC containing WW domains are used for targeted drug screening, i.e. to
CC identify potential modulators of specific WW domain interactions

XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 37; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 28;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
| | | | |
Db 2 ILAPPVP 8

RESULT 11
AAW25513
ID AAW25513 standard; peptide; 13 AA.

XX AAW25513;
AC AAW25513;
XX 27-MAR-1998 (first entry)
DT SH3 synthetic binding peptide.
DE
XX Cortactin; SH3 domain; binding peptide; Src homology region 3;
KW tyrosine kinase; immune response; lymphokine; interleukin 1; Nck; Abl;
KW PLCgamma; p53bp2; Crk; Yes; Grb2.
XX
OS Synthetic.
XX WO9730074-A1.
PN
PD 21-AUG-1997.
XX
PF 14-FEB-1997; 97WO-US002298.
XX
PR 16-FEB-1996; 96US-00602999.
XX
PA (CYTO-) CYTOGEN CORP.
PA (UYNC-) UNIV NORTH CAROLINA.
XX
PI Sparks AB, Kay BK, Thorn JM, Quilliam LA, Der CJ, Fowlkes DM;
PI Rider JE;
XX
DR WPI; 1997-424972/39.
XX

PT Src homology region 3 binding peptide - used to activate Src tyrosine
PT kinase(s) and to stimulate immune response by increasing production of
PT certain lymphokine(s), e.g. interleukin-1.

PS Disclosure; Fig 7; 131pp; English.
XX
CC The present sequence represents a Src SH3 synthetic binding peptide. SH3
CC (Src homology region 3) binding peptides are selected from: (a) peptides
CC which bind the SH3 domain of Cortactin; (b) peptides which bind the
CC middle SH3 domain of Nck; (c) peptides which bind the SH3 domain of Abl;
CC (d) peptides which bind the SH3 domain of Src; (e) peptides which bind
CC the SH3 domain of PLC gamma; (f) peptides which bind the SH3 domain of
CC p53bp2; (g) peptides which bind the amino-terminal SH3 domain of Crk; (h)
CC peptides which bind the SH3 domain of Yes; and (i) peptides which bind
CC the amino-terminal SH3 domain of Grb2. The purified binding peptides can
CC be used in the method to identify inhibitors of their binding to their
CC respective SH3 domains, which could be used to modulate the
CC pharmacological activity of proteins or polypeptide containing the SH3
CC domain. The peptides can also be used to activate Src or Src-related
CC protein tyrosine kinases, to stimulate the immune response by increasing
CC the production of certain lymphokines, e.g. tumour necrosis factor-alpha
CC and interleukin-1, or to deliver a conjugated molecule to certain
CC cellular compartments containing Src or Src related proteins
XX

SQ Sequence 13 AA;

Query Match 100.0%; Score 37; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
| | | | |
Db 2 ILAPPVP 8

RESULT 12
ADB49256
ID ADB49256 standard; peptide; 13 AA.

XX ADB49256;
AC ADB49256;
XX
DT 04-DEC-2003 (first entry)
XX
DE Biotinylated WW domain binding peptide #37.
XX
KW WW domain; drug candidate screening; drug discovery; drug modification;
KW drug refinement; immunogen; WW binding protein; WW domain.
XX
OS Unidentified.
XX US2003077577-A1.
PN
XX
PD 24-APR-2003.
XX
PF 28-JUN-2002; 2002US-00185050.
XX
PR 03-APR-1996; 96US-00630916.
PR 03-APR-1997; 97US-00826516.
XX
PA (PIRO/) PIROZZI G.
PA (KAYB/) KAY B K.
PA (FOWL/) FOWLKES D M.
XX
PI Pirozzi G, Kay BK, Fowlkes DM;
XX WPI; 2003-635075/60.

XX Novel purified polypeptide comprising WW domain, useful for drug
PT discovery, modification and refinement, for discovering polypeptides
PT involved in pharmacological activities, or as an immunogen to generate
PT antibodies.
XX
PS Example; Fig 15A; 133pp; English.
XX
CC The invention describes a purified polypeptide (I) comprising a WW domain
CC which has a sequence (S1) selected from 11 sequences fully defined in the
CC specification, a sequence (S2) selected from 48 sequences fully defined
CC in the specification or a sequence (S3) comprising 683, 906, 224 or 725
CC amino acids fully defined in the specification. (I) is useful for
CC screening a potential drug candidate, by allowing (I) to come into
CC contact with at least one recognition unit having a selective affinity
CC for the WW domain in (I), in the presence of an amount of a potential
CC drug candidate, such that (I) and the recognition unit are capable of
CC interacting when brought into contact with one another in the absence of
CC the drug candidate, and determining the effect, if any, of the presence
CC of the amount of the drug candidate on the interaction of (I) with the
CC recognition unit. (I) is useful for drug discovery, modification and
CC refinement, for discovering polypeptides involved in pharmacological
CC activities, or as an immunogen to generate antibodies. This is the amino
CC acid sequence of a WW domain binding peptide.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 37; DB 7; Length 13;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
| | | | |
Db 2 ILAPPVP 8

RESULT 13
AAW24028
ID AAY24028 standard; peptide; 16 AA.
XX
AC AAY24028;

XX 29-SEP-1999 (first entry)
DT Synthetic peptide comprising a PDZ-binding domain.
XX
DE Human; MMSC1 protein; MMAC1 interacting protein; tumour suppression;
XX MMAC1 pathway; immunogen; cancer; cell neoplastic growth.
KW
KW Synthetic.
OS
OS WO9936566-A1.
PN
XX 22-JUL-1999.
PD
XX 19-JAN-1999; 99WO-US000995.
PF
XX 20-JAN-1998; 98US-0071861P.
PR
XX (MYRI-) MYRIAD GENETICS INC.
PA
XX Bartel PL, Tavtigian SV;
PI
XX WPI; 1999-458472/38.
DR
XX MMSC1, an MMAC1 (tumor suppressor) interacting protein and related
PT polynucleotides.
PT
XX Example 4; Page 48; 107pp; English.
PS
XX The present sequence represents a PDZ-binding domain peptide, used in the
CC course of the invention. The specification describes a MMSC1 protein, a
CC MMAC1 interacting protein involved in tumour suppression activity in the
CC MMAC1 pathway. MMSC1, antigenic fragments or fusion proteins of these are
CC used as immunogens for antibody production. Primers derived from MMSC1
CC genomic clones can be used for identification of MMSC1 genes and for
CC synthesis by amplification of MMSC1 DNA or RNA. Detecting an alteration
CC in MMSC1 can be used to diagnose cancer. A germline alteration in an
CC MMSC1 gene is indicative of a predisposition to cancer. A somatic
CC mutation in an MMSC1 gene is indicative that the tissue is cancerous.
CC Analysis of MMAC1 and MMSC1 (or PDZ domain 6 of MMSC1) binding
CC interactions can be used for detection of alterations in MMAC1 associated
CC with cancer. Wild-type MMSC1 or a homologue can be used to supply wild-
CC type MMSC1 gene function (or a substantially similar function) to a cell,
CC which has lost the gene function due to a MMSC1 gene mutation. The gene
CC suppresses neoplastic growth of the cell. Transgenic animals having an
CC altered MMSC1 can be used as a model for identifying drug candidates
CC useful in treating cancer
XX
SQ Sequence 16 AA;
Query Match 100.0%; Score 37; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ILAPPVP 7
Db 5 ILAPPVP 11
RESULT 14
AAW05412
ID AAW05412 standard; peptide; 17 AA.
XX
AC AAW05412;
XX
DT 24-FEB-1998 (first entry)
XX
DE Src SH3 domain-binding peptide, pSrcCII.
XX
KW Src-homology region 3 domain; human; mouse; SH3 domain; cell growth;
KW cellular signalling element; cellular structural element; malignancy;
KW protein identification; functional domain; protein screening;
KW cellular signal transduction process; binding peptide.

XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH Modified-site 1
FT /note= "Biotin labelled"
FT Modified-site 17
FT /note= "C-terminal amide"
XX
PN WO9631625-A1.
XX
PD 10-OCT-1996.
XX
PF 04-APR-1996; 96WO-US004454.
XX
PR 07-APR-1995; 95US-00417872.
PR 03-APR-1996; 96US-00630915.
XX
PA (CYTO-) CYTOGEN CORP.
PA (UYNC-) UNIV NORTH CAROLINA.
XX
PI Sparks AB, Hoffman N, Kay BK, Fowlkes DM, McConnell SJ;
XX
PI WPI; 1996-465045/46.
DR
XX Identifying polypeptide(s) having specific functional domain (esp. SH3
PT domain) - comprises detecting selective binding to recognition unit,
PT regardless of sequence homology.
XX
PS Example; Page 76; 174pp; English.
XX
CC This sequence represents the Src Src-homology region 3 (SH3) domain-
CC binding peptide termed pSrcCII. pSrcCII was used to identify human and
CC mouse SH3 domain containing proteins (such as AAW05386), that can be used
CC in the method of the invention. The method of the invention is for
CC identifying polypeptides containing functional domains of interest
CC (especially SH3 domains). It comprises contacting a multivalent
CC recognition unit (RU) complex with a number of peptides and identifying
CC polypeptides having a selective binding affinity for the RU complex. The
CC method is based on functional similarities and does not rely on sequence
CC similarities. Prior methods only gave limited success for identifying
CC proteins containing an SH3 domain due to the minimal sequence homology
CC among known SH3 proteins. Multivalent RU complexes are particularly
CC suited to screening for polypeptides containing functional domains that
CC are similar to, but not identical in sequence to, the original target
CC functional domain. The new method enables proteins having a common
CC function to be identified. Identification of novel SH3 proteins will be
CC useful for a better understanding of cell growth, malignancy, signal
CC transduction processes, etc. New candidate drugs can be identified, and
CC their specificities (e.g. pharmacological activities) can be assessed
CC using the method of the invention
XX
SQ Sequence 17 AA;
Query Match 100.0%; Score 37; DB 2; Length 17;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ILAPPVP 7
Db 6 ILAPPVP 12
RESULT 15
AAW16938
ID AAW16938 standard; peptide; 30 AA.
XX
AC AAW16938;
XX
DT 27-JUN-1997 (first entry)
XX
DE Random recombinant SH3 domain binding peptide.
XX

KW Src; SH3; Src homology region 3; binding affinity; oncogenic protein;
KW protein tyrosine kinase; signal transduction; RNA processing;
KW trafficking; translation.
XX
OS Synthetic.
XX
PN WO9603649-A1.
XX
PD 08-FEB-1996.
XX
PF 24-JUL-1995; 95WO-US009382.
XX
PR 22-JUL-1994; 94US-00278865.
PR 07-JUN-1995; 95US-00483555.
XX
PA (UYNC-) UNIV NORTH CAROLINA.
XX
PI Sparks AB, Kay BK, Thorn JM, Quilliam LA, Der CJ;
XX WPI; 1996-117151/12.
XX
PT Peptide with binding affinity for Src homology region 3 (SH3) domains of
PT proteins - useful for e.g. modulating signal transduction pathways at the
PT cellular level, esp. protein tyrosine kinase-mediated.
XX
PS Disclosure; Fig 1; 116pp; English.
XX
CC AAW16924-W16948 are random recombinant peptides derived from one of three
CC peptide libraries, T9, T12 and R8C. The peptides are all SH3 domain-
CC binding peptides. SH3 binding peptides are useful in modulating signal
CC transduction pathways at the cellular level (especially protein tyrosine
CC kinase-mediated), modulating oncogenic protein activity, or providing
CC compounds for the development of drugs with the ability to modulate broad
CC classes, as well as specific classes, of proteins involved in signal
CC transduction and also for regulating the processing, trafficking or
CC translation of RNA. Conjugates of the peptides with detectable labels or
CC imaging agents are useful for imaging cells, tissues and organs in which
CC Src or Src-related proteins are expressed
XX
SQ Sequence 30 AA;

Query Match 100.0%; Score 37; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
Db 16 ILAPPVP 22

RESULT 16
AAW25501
ID AAW25501 standard; peptide; 30 AA.
XX
AC AAW25501;
XX
DT 27-MAR-1998 (first entry)
XX
DE Random peptide recombinant clone T12.SRC3.1.
XX
KW Cortactin; SH3 domain; binding peptide; Src homology region 3;
KW tyrosine kinase; immune response; lymphokine; interleukin 1; Nck; Abl;
KW PLCgamma; p53bp2; Crk; Yes; Grb2.
XX
OS Synthetic.
OS Unidentified.
XX
PN WO9730074-A1.
XX
PD 21-AUG-1997.
XX
PF 14-FEB-1997; 97WO-US002298.
XX

PR 16-FEB-1996; 96US-00602999.
XX
PA (CYTO-) CYTOGEN CORP.
PA (UYNC-) UNIV NORTH CAROLINA.
XX
PI Sparks AB, Kay BK, Thorn JM, Quilliam LA, Der CJ, Fowlkes DM;
PI Rider JE;
XX
DR WPI; 1997-424972/39.
XX
PT Src homology region 3 binding peptide - used to activate Src tyrosine
PT kinase(s) and to stimulate immune response by increasing production of
PT certain lymphokine(s), e.g. interleukin-1.
XX
PS Disclosure; Fig 5; 131pp; English.
XX
CC The present sequence represents a random peptide recombinant isolated by
CC the method of the present invention. SH3 (Src homology region 3) binding
CC peptides are selected from: (a) peptides which bind the SH3 domain of
CC Cortactin; (b) peptides which bind the middle SH3 domain of Nck; (c)
CC peptides which bind the SH3 domain of Abl; (d) peptides which bind the
CC SH3 domain of Src; (e) peptides which bind the SH3 domain of PLC gamma;
CC (f) peptides which bind the SH3 domain of p53bp2; (g) peptides which bind
CC the amino-terminal SH3 domain of Crk; (h) peptides which bind the SH3
CC domain of Yes; and (i) peptides which bind the amino-terminal SH3 domain
CC of Grb2. The purified binding peptides can be used in the method to
CC identify inhibitors of their binding to their respective SH3 domains,
CC which could be used to modulate the pharmacological activity of proteins
CC or polypeptide containing the SH3 domain. The peptides can also be used
CC to activate Src or Src-related protein tyrosine kinases, to stimulate the
CC immune response by increasing the production of certain lymphokines, e.g.
CC tumour necrosis factor-alpha and interleukin-1, or to deliver a
CC conjugated molecule to certain cellular compartments containing Src or
CC Src related proteins
XX
SQ Sequence 30 AA;

Query Match 100.0%; Score 37; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
Db 16 ILAPPVP 22

RESULT 17
ABB62726
ID ABB62726 standard; protein; 1767 AA.
XX
AC ABB62726;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 14970.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.
DR N-PSDB; ABL06829.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Disclosure; SEQ ID NO 14970; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1767 AA;

Query Match 100.0%; Score 37; DB 4; Length 1767;
Best Local Similarity 100.0%; Pred. No. 2.8e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
Db 1101 ILAPPVP 1107

RESULT 18
AAW25387
ID AAW25387 standard; peptide; 16 AA.
XX
AC AAW25387;
XX
DT 27-MAR-1998 (first entry)
XX
DE PLC gamma SH3 domain binding peptide SEQ ID NO:178.
XX
KW Cortactin; SH3 domain; binding peptide; Src homology region 3;
KW tyrosine kinase; immune response; lymphokine; interleukin 1; Nck; Abl;
KW PLCgamma; p53bp2; Crk; Yes; Grb2.
XX
OS Synthetic.
OS Unidentified.
XX
PN WO9730074-A1.
XX
PD 21-AUG-1997.
XX
PF 14-FEB-1997; 97WO-US002298.
XX
PR 16-FEB-1996; 96US-00602999.
XX
PA (CYTO-) CYTOGEN CORP.
PA (UYNC-) UNIV NORTH CAROLINA.
XX
PI Sparks AB, Kay BK, Thorn JM, Quilliam LA, Der CJ, Fowlkes DM;
PI Rider JE;
XX
DR WPI; 1997-424972/39.
XX
PT Src homology region 3 binding peptide - used to activate Src tyrosine
PT kinase(s) and to stimulate immune response by increasing production of
PT certain lymphokine(s), e.g. interleukin-1.
XX
PS Claim 15; Page 98; 131pp; English.
XX
CC The present sequence represents a Src homology region 3 (SH3) binding
CC peptide. SH3 binding peptides are selected from: (a) peptides which bind

CC the SH3 domain of Cortactin; (b) peptides which bind the middle SH3
CC domain of Nck; (c) peptides which bind the SH3 domain of Abl; (d)
CC peptides which bind the SH3 domain of Src; (e) peptides which bind the
CC SH3 domain of PLC gamma; (f) peptides which bind the SH3 domain of p53bp2
CC ; (g) peptides which bind the amino-terminal SH3 domain of Crk; (h)
CC peptides which bind the SH3 domain of Yes; and (i) peptides which bind
CC the amino-terminal SH3 domain of Grb2. The purified binding peptides can
CC be used in the method to identify inhibitors of their binding to their
CC respective SH3 domains, which could be used to modulate the
CC pharmacological activity of proteins or polypeptide containing the SH3
CC domain. The peptides can also be used to activate Src or Src-related
CC protein tyrosine kinases, to stimulate the immune response by increasing
CC the production of certain lymphokines, e.g. tumour necrosis factor-alpha
CC and interleukin-1, or to deliver a conjugated molecule to certain
CC cellular compartments containing Src or Src related proteins
XX
SQ Sequence 16 AA;

Query Match 97.3%; Score 36; DB 2; Length 16;
Best Local Similarity 85.7%; Pred. No. 49;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
Db 1 VLAPPVP 7

RESULT 19
ABP43530
ID ABP43530 standard; protein; 137 AA.
XX
AC ABP43530;
XX
DT 08-AUG-2002 (first entry)
XX
DE Human secreted protein (SCEP) 54.
XX
KW Human; secreted protein; SECP; SECP expression; gene therapy;
KW protein therapy; immune system disorders; AIDS; thymic hypoplasia;
KW anaemia; asthma; Crohn's disease; neurological disorder; epilepsy;
KW Huntington's disease; dementia; Parkinson's disease; Down's syndrome;
KW developmental disorder; cell proliferative disorder; cancer.
XX
OS Homo sapiens.
XX WO200226982-A2.
XX
PD 04-APR-2002.
XX
PF 25-SEP-2001; 2001WO-US030042.
XX
PR 29-SEP-2000; 2000US-0236869P.
PR 11-OCT-2000; 2000US-0239812P.
PR 12-OCT-2000; 2000US-0240108P.
PR 17-OCT-2000; 2000US-0241282P.
PR 20-OCT-2000; 2000US-0242218P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Yue H, Tang YT, Nguyen DB, Yao MG, Xu Y, Tribouley CM;
PI Sanjanwala MS, Wallia NK, Baughn MR, Sapperstein SK, Lal P;
PI Thornton M, Gandhi AR, Ramkumar J, Elliott VS, Arvizu C;
PI Thangavelu K, Gietzen KJ, Ding L, Au-Young J, Tran B, Policky JL;
PI Lee S, Lu DAM, Burford N, Warren BA, Gururajan R, Duggan BM;
PI Honchell CD, Hafalia AJA;
XX
DR WPI; 2002-394239/42.
DR N-PSDB; ABN99413.
XX
PT New human secreted proteins, useful for diagnosing, treating or
PT preventing immune system disorders (e.g. Crohn's disease), neurological
PT disorders (e.g. Parkinson's disease), or cell proliferative disorders
PT (e.g. cancers).

XX PS Claim 1; Page 186; 238pp; English.

XX CC The invention comprises the amino acid and coding sequences of human

CC secreted proteins (SECP). The SECP DNA and amino acid sequences of the

CC invention are useful for treating/preventing disorders associated with

CC decreased or elevated expression of SECP. The SECP DNA and protein

CC sequences are specifically useful for treating/preventing (i.e. gene

CC therapy and protein therapy): immune system disorders (e.g. AIDS, thymic

CC hypoplasia, anaemia, asthma or Crohn's disease); neurological disorders

CC (e.g. epilepsy, Huntington's disease, dementia or Parkinson's disease);

CC developmental disorders (e.g. Down's syndrome); and cell proliferative

CC disorders (e.g. cancer). The proteins ABP43477 - ABP43543 represent the

CC human secreted proteins (SECP) of the invention

XX SQ Sequence 137 AA;

Query Match 97.3%; Score 36; DB 5; Length 137;

Best Local Similarity 85.7%; Pred. No. 3.7e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7

Db 15 VLAPPVP 21

RESULT 20

AAG83323

ID AAG83323 standard; protein; 159 AA.

XX AC AAG83323;

XX DT 06-SEP-2001 (first entry)

XX DE P patens lipid metabolism related protein #41.

XX KW Moss; LMRP; lipid metabolism related protein; polyunsaturated fatty acid;

KW fine chemical; transgenic plant.

XX OS Physcomitrella patens.

XX PN WO200138541-A1.

XX PD 31-MAY-2001.

XX PF 25-NOV-1999; 99WO-EP009108.

XX PR 25-NOV-1999; 99WO-EP009108.

XX PA (BADI) BASF PLANT SCI GMBH.

XX PI Lerchl J, Renz A, Ehrhardt T, Reindl A, Cirpus P, Bischoff F;

PI Frank M, Freund A, Duwenig E, Schmidt R, Reski R;

XX DR WPI; 2001-381293/40.

XX PT New isolated nucleic acid molecule encoding Lipid Metabolism Related

PT Proteins useful in the production of fine chemicals.

XX PS Claim 31; Page 103; 113pp; English.

XX CC The present invention provides the protein and coding sequences of a

CC number of moss lipid metabolism related proteins (LMRPs). The moss

CC Physcomitrella patens is one of the few plants able to produce

CC polyunsaturated fatty acids, and the sequences can be used to create

CC transgenic plants also capable of producing them. They can also be used

CC to identify the presence of p. patens and in the production of fine

CC chemicals. The present sequence is one of the proteins of the invention

XX SQ Sequence 159 AA;

Query Match 97.3%; Score 36; DB 4; Length 159;

Best Local Similarity 85.7%; Pred. No. 4.2e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7

Db 41 VLAPPVP 47

RESULT 21

AAG80875

ID AAG80875 standard; protein; 159 AA.

XX AC AAG80875;

XX DT 28-AUG-2001 (first entry)

XX DE Lipid biosynthesis protein sequence #33.

XX KW Moss; Physcomitrella patens; lipid metabolism related protein; LMRP;

KW lipid biosynthesis; lipid modification; lipid degradation; cofactor;

KW fatty acid transport; genetic engineering; fatty acid; enzyme; plant;

KW microorganism; polyunsaturated fatty acid; oilseed plant; maize; wheat;

KW biotic stress tolerance; abiotic stress tolerance; rye; oat; triticale;

KW rice; barley; soybean; peanut; cotton; rapeseed; canola; manihot; pepper;

KW sunflower; tagetes; potato; tobacco; eggplant; tomato; Vicia; pea;

KW alfalfa; coffee; cacao; tea; Salix; oil palm; coconut; perennial grass;

KW forage crop.

XX OS Physcomitrella patens.

XX PN WO200138484-A2.

XX PD 31-MAY-2001.

XX PF 22-NOV-2000; 2000WO-EP011615.

XX PR 25-NOV-1999; 99WO-EP009108.

XX PA (BADI) BASF PLANT SCI GMBH.

XX PI Lerchl J, Renz A, Ehrhardt T, Reindl A, Cirpus P, Bischoff F;

PI Frank M, Freund A, Duwenig E, Schmidt R, Reski R;

XX DR WPI; 2001-367669/38.

XX PT Nucleic acids encoding lipid metabolism related proteins from

PT Physcomitrella patens useful to produce fine chemicals in modified

PT organisms, particularly polyunsaturated fatty acids in oilseed plants.

XX PS Claim 31; Page 113; 120pp; English.

XX CC The present invention describes isolated nucleic acid sequences which

CC encode lipid metabolism related proteins (LMRP). The LMRP nucleic acids

CC can be used to modify lipids and fatty acids, cofactors and enzymes in

CC microorganisms and plants, particularly to produce polyunsaturated fatty

CC acids, and are especially useful in oilseed plants. The nucleic acids may

CC also confer biotic or abiotic stress tolerance, particularly to maize,

CC wheat, rye, oat, triticale, rice, barley, soybean, peanut, cotton,

CC rapeseed, canola, manihot, pepper, sunflower, tagetes, potato, tobacco,

CC eggplant, tomato, Vicia species, pea, alfalfa, coffee, cacao, tea, Salix

CC species, oil palm, coconut, perennial grasses and forage crops. AAH50878

CC to AAH50882 represent primers used in the exemplification of the present

CC invention. AAH50883 to AAH50968 represents LMRP nucleotide sequences, and

CC AAG80843 to AAG80928 represent LMRP protein sequences, given in the

CC present invention

XX SQ Sequence 159 AA;

Query Match 97.3%; Score 36; DB 4; Length 159;

Best Local Similarity 85.7%; Pred. No. 4.2e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7

Db 41 VLAPPVP 47

Db 41 VLAPPVP 47

RESULT 22
ABB50640
ID ABB50640 standard; protein; 200 AA.
XX
AC ABB50640;
XX
DT 07-FEB-2002 (first entry)
XX
DE Human secreted protein encoded by gene 39 SEQ ID NO:588.
XX
KW Human; secreted protein; immunomodulatory; antisclerotic; anti-HIV;
KW dermatological; immunosuppressive; antiinflammatory; immunostimulant;
KW cytosolic; cardiant; vascular; anti-angiogenic; ophthalmological;
KW neuroprotective; nootropic; anticonvulsant; antialzheimers; vulnerary;
KW antiparkinsonian; antimicrobial; gene therapy; vaccine; immune disorder;
KW multiple sclerosis; systemic lupus erythematosus; HIV infection; cancer;
KW human immunodeficiency virus; hyperproliferative disorder; wound healing;
KW Gaucher's disease; cardiovascular disease; Scimitar syndrome; chemotaxis;
KW Chaga's cardiomyopathy; coronary arteriosclerosis; angiogenic disorder;
KW corneal graft neovascularisation; diabetic retinopathy; regeneration;
KW neurological disorder; Huntington's chorea; Alzheimer's disease;
KW Parkinson's disease; infectious disease.
XX
OS Homo sapiens.
XX
PN WO200162891-A2.
XX
PD 30-AUG-2001.
XX
PF 21-FEB-2001; 2001WO-US005614.
XX
PR 24-FEB-2000; 2000US-0184836P.
PR 29-MAR-2000; 2000US-0193170P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ni J, Ebner R, Lafleur DW, Moore PA, Olsen HS, Rosen CA;
PI Ruben SM, Soppet DR, Young PE, Shi Y, Florence KA, Wei Y;
PI Florence C, Hu J, Li Y, Kyaw H, Fischer CL, Ferrie AM, Fan P;
PI Feng P, Endress GA, Dillon PJ, Carter KC, Brewer LA, Yu G, Zeng Z;
PI Greene JM;
XX
DR WPI; 2001-625724/72.
XX
XX
PT Nucleic acids encoding 207 human secreted polypeptides, useful for
PT preventing, diagnosing and/or treating, e.g. cancers, Parkinson's disease
PT and diabetic retinopathy.
XX
PS Disclosure; Page 90; 1533pp; English.
XX
CC ABB50301 to ABB51287 and ABA83194 to ABA83441 represent human secreted
CC proteins (I) and polynucleotide (II) sequences. (I) and (II) have various
CC activities based on the tissues and cells the genes are expressed in.
CC Example of these activities include: immunomodulatory; antisclerotic;
CC dermatological; immunosuppressive; antiinflammatory; immunostimulant;
CC anti-HIV; cytostatic; cardiant; anti-angiogenic; ophthalmological;
CC neuroprotective; nootropic; anticonvulsant; antialzheimers; vascular;
CC antiparkinsonian; antimicrobial; and vulnerary. (I) and (II) can be used
CC in gene therapy and vaccine production. (I) and (II) can be used in the
CC prevention, diagnosis and treatment of immune disorders (e.g. multiple
CC sclerosis, systemic lupus erythematosus and human immunodeficiency virus
CC (HIV) infections), hyperproliferative disorders (e.g. cancers and
CC Gaucher's disease), cardiovascular diseases (e.g. Scimitar syndrome,
CC Chaga's cardiomyopathy and coronary arteriosclerosis), angiogenic
CC disorders (e.g. corneal graft neovascularisation and diabetic
CC retinopathy), neurological disorders (e.g. Huntington's chorea,
CC Alzheimer's disease and Parkinson's disease), infectious diseases and/or
CC for promoting wound healing, regeneration and/or chemotaxis. ABA83185 to
CC ABA83193 and ABB50300 represent sequences used in the exemplification of
CC the present invention

XX SQ Sequence 200 AA;
Query Match 97.3%; Score 36; DB 4; Length 200;
Best Local Similarity 85.7%; Pred. No. 5.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 ILAPPVP 7
Db 160 VLAPPVP 166
RESULT 23
ABO44897
ID ABO44897 standard; protein; 200 AA.
XX
AC ABO44897;
XX
DT 02-OCT-2003 (first entry)
XX
DE Novel human secreted protein #39 fragment #3.
XX
KW Human; gene therapy; autoimmune disorder; multiple sclerosis; cancer;
KW systemic lupus erythematosus; haematopoietic cell disorder; allergy;
KW agammaglobulinaemia; ataxia telangiectasia; blood coagulation disorder;
KW afibrinogenaemia; thrombocytopenia; graft-versus-host disease; arthritis;
KW inflammatory condition; ischaemia-reperfusion injury; infectious disease;
KW hyperproliferative disorder; purpura; viral infection; regeneration;
KW bacterial infection; ulcer; Alzheimer's disease.
XX
OS Homo sapiens.
XX
PN US2003065160-A1.
XX
PD 03-APR-2003.
XX
PF 07-DEC-2001; 2001US-00004860.
XX
PR 06-JUN-1997; 97US-0048875P.
PR 06-JUN-1997; 97US-0048876P.
PR 06-JUN-1997; 97US-0048877P.
PR 06-JUN-1997; 97US-0048878P.
PR 06-JUN-1997; 97US-0048880P.
PR 06-JUN-1997; 97US-0048881P.
PR 06-JUN-1997; 97US-0048882P.
PR 06-JUN-1997; 97US-0048883P.
PR 06-JUN-1997; 97US-0048884P.
PR 06-JUN-1997; 97US-0048885P.
PR 06-JUN-1997; 97US-0048892P.
PR 06-JUN-1997; 97US-0048893P.
PR 06-JUN-1997; 97US-0048894P.
PR 06-JUN-1997; 97US-0048895P.
PR 06-JUN-1997; 97US-0048896P.
PR 06-JUN-1997; 97US-0048897P.
PR 06-JUN-1997; 97US-0048898P.
PR 06-JUN-1997; 97US-0048899P.
PR 06-JUN-1997; 97US-0048900P.
PR 06-JUN-1997; 97US-0048901P.
PR 06-JUN-1997; 97US-0048915P.
PR 06-JUN-1997; 97US-0048916P.
PR 06-JUN-1997; 97US-0048917P.
PR 06-JUN-1997; 97US-0048949P.
PR 06-JUN-1997; 97US-0048962P.
PR 06-JUN-1997; 97US-0048963P.
PR 06-JUN-1997; 97US-0048964P.
PR 06-JUN-1997; 97US-0048970P.
PR 06-JUN-1997; 97US-0048971P.
PR 06-JUN-1997; 97US-0048972P.
PR 06-JUN-1997; 97US-0048974P.
PR 06-JUN-1997; 97US-0049019P.
PR 06-JUN-1997; 97US-0049020P.
PR 06-JUN-1997; 97US-0049373P.
PR 06-JUN-1997; 97US-0049374P.

PR 06-JUN-1997; 97US-0049375P.
PR 05-SEP-1997; 97US-0057584P.
PR 05-SEP-1997; 97US-0057627P.
PR 05-SEP-1997; 97US-0057628P.
PR 05-SEP-1997; 97US-0057629P.
PR 05-SEP-1997; 97US-0057634P.
PR 05-SEP-1997; 97US-0057635P.
PR 05-SEP-1997; 97US-0057642P.
PR 05-SEP-1997; 97US-0057643P.
PR 05-SEP-1997; 97US-0057644P.
PR 05-SEP-1997; 97US-0057645P.
PR 05-SEP-1997; 97US-0057646P.
PR 05-SEP-1997; 97US-0057647P.
PR 05-SEP-1997; 97US-0057648P.
PR 05-SEP-1997; 97US-0057649P.
PR 05-SEP-1997; 97US-0057650P.
PR 05-SEP-1997; 97US-0057651P.
PR 05-SEP-1997; 97US-0057654P.
PR 05-SEP-1997; 97US-0057661P.
PR 05-SEP-1997; 97US-0057662P.
PR 05-SEP-1997; 97US-0057666P.
PR 05-SEP-1997; 97US-0057667P.
PR 05-SEP-1997; 97US-0057668P.
PR 05-SEP-1997; 97US-0057760P.
PR 05-SEP-1997; 97US-0057761P.
PR 05-SEP-1997; 97US-0057762P.
PR 05-SEP-1997; 97US-0057763P.
PR 05-SEP-1997; 97US-0057764P.
PR 05-SEP-1997; 97US-0057765P.
PR 05-SEP-1997; 97US-0057769P.
PR 05-SEP-1997; 97US-0057770P.
PR 05-SEP-1997; 97US-0057771P.
PR 05-SEP-1997; 97US-0057774P.
PR 05-SEP-1997; 97US-0057775P.
PR 05-SEP-1997; 97US-0057776P.
PR 05-SEP-1997; 97US-0057777P.
PR 05-SEP-1997; 97US-0057778P.
PR 18-DEC-1997; 97US-0070923P.
PR 04-JUN-1998; 98WO-US011422.
PR 15-JUL-1998; 98US-0092921P.
PR 30-JUL-1998; 98US-0094657P.
XX (HUMA-) HUMAN GENOME SCI INC.
PI Young P, Greene JM, Ferrie AM, Ruben SM, Rosen CA, Hu J;
PI Olsen HS, Ebner R, Brewer LA, Moore PA, Shi Y, Florence C;
PI Florence K, Lafleur DW, Ni J, Fan P, Wei Y, Fischer CL, Soppet DR;
PI Li Y, Zeng Z, Kyaw H, Yu G, Feng P, Dillon PJ, Endress GA;
PI Carter KC;
XX WPI; 2003-511926/48.
XX New precerebellin-like protein, useful for diagnosing or treating
PT neurodegenerative and behavioral disorders, immune disorders, liver
PT disorders, and cancer.
XX Disclosure; Col 42; 156pp; English.
XX The invention relates to an isolated protein comprising amino acid
CC residues 33-205 or 1-205 of a novel human secreted protein appearing as
CC ABO26252. The protein is encoded by one of 238 disclosed cDNA sequences
CC encoding 238 secreted proteins. ABO26252 is a precerebellin-like protein.
CC Also included are a composition comprising the protein and a carrier and
CC an isolated protein produced by expressing the protein cited above by a
CC cell, and recovering the protein. The proteins are useful for diagnosing
CC or treating neurodegenerative and behavioural disorders (e.g. Alzheimer's
CC disease, Parkinson's disease, Huntington's disease, schizophrenia, mania,
CC dementia, paranoia, psychoses or autism), immune disorders (e.g.
CC infection, inflammation, allergy), liver disorders (e.g. hepatoblastoma,
CC jaundice, hepatitis), immunological disorders (e.g. AIDS, leukaemia,
CC rheumatoid arthritis, sepsis, acne, psoriasis) and cancer. The present
CC sequence is a protein associated with one of the 238 disclosed novel
CC secreted proteins

XX SQ Sequence 200 AA;
Query Match 97.3%; Score 36; DB 7; Length 200;
Best Local Similarity 85.7%; Pred. No. 5.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 ILAPPVP 7
Db 160 VLAPPVP 166
RESULT 25
ADQ67709
ID ADQ67709 standard; protein; 218 AA.
XX
AC ADQ67709;
XX 07-OCT-2004 (first entry)
XX
DE Novel human protein sequence #2375.
XX
KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;
KW gene therapy; diagnostic marker; morbid state; osteoporosis;
KW neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
KW cancer.
XX Homo sapiens.
XX
PN EP1440981-A2.
XX
PD 28-JUL-2004.
XX
PF 21-JAN-2004; 2004EP-00001196.
XX
PR 21-JAN-2003; 2003JP-00102206.
PR 09-MAY-2003; 2003JP-00131392.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Nagai K, Irie R;
XX
DR WPI; 2004-535376/52.
DR N-PSDB; ADQ67402.
XX
PT Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX
PS Claim 1; SEQ ID NO 4870; 2449pp; English.
XX
CC The invention relates to 2495 novel polynucleotides (I) and their encoded
CC polypeptides, sequences hybridizing to these nucleotides, sequences
CC encoding partial polypeptides and sequences having 70% or 90% identity to
CC the nucleotide and protein sequences. The nucleotides and polypeptides
CC are useful as diagnostic markers or therapeutic target for the diseases
CC or morbid states. They are also useful for treating osteoporosis,
CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,
CC dementia and various cancers. This sequence corresponds to a protein
CC sequence of the invention.
XX
SQ Sequence 218 AA;
Query Match 97.3%; Score 36; DB 8; Length 218;
Best Local Similarity 85.7%; Pred. No. 5.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 ILAPPVP 7
Db 93 VLAPPVP 99
RESULT 26

AAU17916
ID AAU17916 standard; protein; 116 AA.
XX
AC AAU17916;
XX
DT 07-NOV-2001 (first entry)
XX
DE Novel human respiratory antigen #232.
XX
KW Human; respiratory antigen; respiratory disorder; throat disorder;
lung disorder; nose disorder; lung cancer; gene therapy; cytostatic;
anti allergic; anti asthmatic; anti inflammatory; olfactory;
respiratory active.
XX
OS Homo sapiens.
XX
PN WO20015448-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001333.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.

PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.

PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-476224/51.
DR N-PSDB; AAS28100.
XX
PT Isolated polypeptide for treating, preventing and/ or prognosing
PT disorders related to the respiratory system including respiratory cancers
PT and also for testing and detection e.g. diagnosis.
XX
PS Claim 11; SED ID No 534; 546pp; English.
XX
CC The present invention relates to the isolation of novel human respiratory
CC antigens, and cDNA (AAS27869-AAS28159) and genomic sequences encoding for
CC these polypeptides. The sequences of the invention are useful for
CC preventing, treating and/or prognosing disorders related to the
CC respiratory system including throat disorders (e.g. vocal cord paralysis,
CC tonsillitis, and laryngitis), lung disorders e.g. pneumonia, allergic
CC disorders e.g. asthma, pleurisy, cystic fibrosis, emphysema, nose
CC disorders and cancers of the respiratory tissues e.g. lung cancer. The
CC polynucleotide sequences of the invention are useful in gene therapy and
CC antisense therapy. AAU1685-AAU1975 represent novel human respiratory
CC antigens. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 116 AA;

Query Match 94.6%; Score 35; DB 4; Length 116;
Best Local Similarity 85.7%; Pred. No. 4.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
Db 77 LLAPPVP 83
:
:|||||

RESULT 27
ADG41296
ID ADG41296 standard; protein; 116 AA.
XX
AC ADG41296;
XX
DT 26-FEB-2004 (first entry)
XX
DE Human respiratory system associated protein seq id 534.
XX
KW antiinflammatory; antiallergic; antiasthmatic; cytostatic; gene therapy;
KW respiratory system antigen;
KW human respiratory system associated polynucleotide;
KW respiratory system disorder; throat disorder; vocal cord paralysis;
KW tonsillitis; laryngitis; lung disorder; pneumonia; allergic disorder;
KW asthma; eosinophilic pneumonia; pleurisy; cystic fibrosis; emphysema;
KW histiocytosis; sarcoidosis; nose disorder; rhinitis; sinusitis; neoplasm;
KW cancer; respiratory tissue cancer; throat cancer; lung cancer;
KW cancer of the nose; gene therapy; chromosome identification; forensic;
KW human respiratory system associated protein; human.
XX
OS Homo sapiens.

XX US2003215893-A1.
PN
XX
PD 20-NOV-2003.
XX
PF 07-AUG-2002; 2002US-00212872.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR

PR 27-SEP-2000; 2000US-0235836P.
PR 28-SEP-2000; 2000US-0235935P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764860.
PR 14-FEB-2002; 2002US-00074095.
XX

PA (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Ruben SM, Barash SC;
PI
XX
DR WPI; 2003-902033/82.
DR N-PSDB; ADG41004.
XX
PT Novel respiratory system antigen and polynucleotides encoding the
PT polypeptides, useful for treating diagnosing, treating or preventing
PT tonsillitis, pneumonia, asthma and cystic fibrosis, emphysema, throat
PT cancer.
XX
PS Claim 11; SEQ ID NO 534; 236pp; English.
XX
CC The invention describes an isolated polypeptide (I) comprising an amino
CC acid sequence that is at least 90% identical to polypeptide fragment of
CC any one of 299 respiratory system antigen sequences (PS) and having
CC biological activity, polypeptide domain or epitope of PS, full-length
CC protein of PS, or variant, allelic variant or species homolog of PS. (I)
CC or a polynucleotide (II) encoding (I) is also useful for diagnosing a
CC pathological condition or a susceptibility to a pathological condition in
CC a subject which involves determining the presence or absence of mutation
CC in (II) or determining the presence or amount of expression of (I) in a
CC biological sample and diagnosing a pathological condition based on the
CC result. The human respiratory system associated polynucleotides, the
CC polypeptides encoded by them, and antibodies that immunospecifically bind
CC these polypeptides are useful in diagnosis, treatment, prevention and/or
CC prognosis of disorders of respiratory system such as throat disorders
CC (e.g., vocal cord paralysis, tonsillitis, and laryngitis), lung disorders
CC (e.g., pneumonia), allergic disorders, (e.g., asthma and eosinophilic
CC pneumonia), pleurisy, cystic fibrosis, emphysema, histiocytosis,
CC sarcoidosis, nose disorders (rhinitis and sinusitis), neoplasms and/or
CC cancers of respiratory tissues (e.g., throat cancer, lung cancer, and
CC cancer of the nose). The polynucleotides are useful in gene therapy

Query Match 94.6%; Score 35; DB 7; Length 116;
Best Local Similarity 85.7%; Pred. No. 4.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
:|||||
Db 77 LLAPPVP 83

RESULT 28
ADI97070
ID ADI97070 standard; protein; 116 AA.
XX
AC ADI97070;
XX
DT 04-NOV-2004 (first entry)
XX
DE Human respiratory system associated polypeptide SeqID534.
XX
KW respiratory system-related polypeptide; antiasthmatic; antibacterial;
KW antiinflammatory; cytostatic; antianaemic; antiallergic; gene therapy;
KW pneumonia; lung cancer; cystic fibrosis; asthma; sarcoidosis; rhinitis;
KW anaemia; leukaemia; inflammation; sinusitis;
KW chronic obstructive pulmonary disease; infectious disease; human.
XX
OS Homo sapiens.
XX
PN US2003077704-A1.
XX
PD 24-APR-2003.
XX
PF 14-FEB-2002; 2002US-00074095.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.

PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764860.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Ruben SM, Barash SC;

WPI; 2003-765403/72.
N-PSDB; ADI96778.

New human respiratory system-related polypeptide and genes, useful for treating, preventing or diagnosing e.g. pneumonia, lung cancer, cystic fibrosis, asthma, sarcoidosis, rhinitis, leukemia, inflammations or sinusitis.

Claim 11; SEQ ID NO 534; 202pp; English.

XX This invention is related to a novel isolated polypeptide, which
CC comprises a human respiratory system-related polypeptide, and the DNA
CC sequence which encodes it. The invention may be useful for the
CC development of compounds with an antiasthmatic, antibacterial,
CC antiinflammatory, cytostatic, antianaemic or antiallergic activity. In
CC addition, the sequences disclosed may be useful for gene therapy. The
CC polypeptide or polynucleotide is useful for treating, preventing or
CC ameliorating a medical condition, for example pneumonia, lung cancer,
CC cystic fibrosis, asthma, sarcoidosis, rhinitis, anaemia, leukaemia,
CC inflammations, sinusitis, chronic obstructive pulmonary disease or
CC infectious diseases. The polypeptide or polynucleotide is also useful for
CC diagnosing any of these diseases or a susceptibility to the disease. The
CC present sequence is that of a human respiratory system associated
XX polypeptide of the invention.

SQ Sequence 116 AA;

Query Match 94.6%; Score 35; DB 7; Length 116;
Best Local Similarity 85.7%; Pred. No. 4.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
Db :|||||
77 LLAPPVP 83

RESULT 29
AAG10732
ID AAG10732 standard; protein; 132 AA.
XX
AC AAG10732;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 9171.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.

PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 22-OCT-1999; 99US-0160989P.

PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 94.6%; Score 35; DB 3; Length 132;
Best Local Similarity 85.7%; Pred. No. 5.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
:|||||
Db 102 LLAPPVP 108

RESULT 30
ABM90479
ID ABM90479 standard; protein; 477 AA.
XX
AC ABM90479;
XX
DT 02-JUN-2005 (first entry)
XX
DE Rice abiotic stress responsive polypeptide SEQ ID NO:9201.
XX
KW abiotic stress tolerance; transgenic plant; plant; cereal; agriculture.
XX
OS Oryza sativa.
XX
PN WO2003008540-A2.
XX
PD 30-JAN-2003.
XX
PF 21-JUN-2002; 2002WO-US019668.
XX
PR 22-JUN-2001; 2001US-0300112P.
PR 24-AUG-2001; 2001US-0314662P.
PR 26-SEP-2001; 2001US-0325277P.
PR 21-NOV-2001; 2001US-0332132P.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Kreps J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F;
PI Moughamer T, Provart N, Ricke D, Zhu T;
XX
DR WPI; 2003-248011/24.
XX
PT New stress-responsive nucleic acid, useful for altering the
PT responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
PT stress, salt stress or osmotic stress.
XX
PS Claim 1; SEQ ID NO 9201; 89pp; English.
XX
CC The invention relates to novel abiotic stress responsive polynucleotides
CC and polypeptides. Also disclosed are vectors, expression cassettes, host
CC cells, and plants containing such polynucleotides. Also disclosed are
CC methods for using the polynucleotides and polypeptides to alter the
CC responsiveness of a plant to abiotic stress. The invention is useful in
CC agriculture. The nucleic acid is useful for determining whether a test
CC plant has been exposed to an abiotic stress condition. It is also useful
CC for selecting an agent that alters abiotic stress regulated
CC polynucleotide expression in a plant cell, and to identify a homolog or
CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
CC molecule and the polypeptide encoded by it are useful in altering the
CC responsiveness of a plant to an abiotic stress, such as cold stress, salt
CC stress, osmotic stress or any of their combinations. The present sequence
CC is used in the exemplification of the invention
XX

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
||:||||
Db 70 ILSPPVP 76

RESULT 33
ADC97354
ID ADC97354 standard; protein; 177 AA.

XX AC ADC97354;

XX DT 01-JAN-2004 (first entry)

XX DE E. faecium protein sequence SEQ ID 6981.

XX KW Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;
XX KW abdominal-pelvic infection.

XX OS Enterococcus faecium.

XX PN U96583275-B1.

XX PD 24-JUN-2003.

XX PF 30-JUN-1998; 98US-00107532.

XX PR 02-JUL-1997; 97US-0051571P.

XX PR 14-MAY-1998; 98US-0085598P.

XX PA (GENO-) GENOME THERAPEUTICS CORP.

XX PI Doucette-Stamm LA, Bush D;

XX DR WPI; 2003-799836/75.

XX DR N-PSDB; ADC93700.

XX PT New isolated nucleic acid derived from Enterococcus faecium encoding an
XX PT Enterococcus faecium polypeptide useful for detection, prevention and
XX PT treatment of a pathological condition resulting from a bacterial
XX PT infection.

XX PS Example 1; SEQ ID NO 6981; 243pp; English.

XX CC The invention relates to an isolated nucleic acid derived from
XX CC Enterococcus faecium encoding an Enterococcus faecium polypeptide having
XX CC one of 10 fully defined sequences given in the (or comprising 40
XX CC sequential nucleotides chosen from any of the nucleic acids, its
XX CC complement or sequences hybridising to it). Also included are a
XX CC recombinant vector comprising the nucleic acid operably linked to
XX CC transcription regulatory element, a cell comprising the vector and a
XX CC single-stranded probe comprising the nucleic acid. The nucleic acids are
XX CC chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.
XX CC The nucleic acids is useful for diagnosing pathological conditions
XX CC resulting from E. faecium bacterial infection (e.g. urinary tract
XX CC infection, bacteraemia, endocarditis, wounds and abdominal-pelvic
XX CC infection) and for screening drugs such as agonists and antagonists. The
XX CC nucleic acid is useful for recombinant production of Candida albicans -
XX CC derived peptides or antisense polypeptides. Pharmaceutical compositions
XX CC and vaccines containing the nucleic acid are useful for preventing or
XX CC treating Enterococcus faecium infections. The present sequence represents
XX CC one if the disclosed E. faecium proteins.

XX SQ Sequence 177 AA;

Query Match 91.9%; Score 34; DB 7; Length 177;
Best Local Similarity 71.4%; Pred. No. 9.6e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
||:||||
Db 63 IIAPPPI 69

RESULT 34
ADK34392
ID ADK34392 standard; protein; 291 AA.

XX AC ADK34392;

XX DT 06-MAY-2004 (first entry)

XX DE Novel human polypeptide SeqID6474.

XX KW antiarthritic; antiparkinsonian; neuroprotective; nootropic;
XX KW immunosuppressive; cytostatic; antipsoriatic; antiinflammatory;
XX KW antibacterial; antiviral; antifungal; antiparasitic; gene therapy;
XX KW arthritis; Parkinson's; Alzheimer's; autoimmune disease; cancer;
XX KW psoriasis; inflammatory bowel disease; infection; bacteria; virus;
XX KW fungus; parasite; human.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Misc-difference 1..291

FT /label= OTHER

FT /note= "OTHER= All Xaa's in this sequence are unknown

FT amino acids or the site of a stop codon within the DNA

FT sequence"

XX WO200216439-A2.

XX PN 28-FEB-2002.

XX PD 05-MAR-2001; 2001WO-US004941.

XX PF 07-MAR-2000; 2000US-00519705.

XX PR 19-MAY-2000; 2000US-00574454.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX DR WPI; 2002-280918/32.

XX PS Isolated polynucleotide encoding bone marrow derived polypeptides useful
XX PS for treating, e.g., Parkinson's, Alzheimer's, cancer, arthritis, Crohn's
XX PS disease, and inflammatory bowel disease.

XX PS Claim 20; SEQ ID NO 6474; 504pp; English.

XX CC This invention relates to a novel isolated polynucleotide comprising a
XX CC nucleotide sequence selected from one of 1680 sequences, a mature protein
XX CC coding portion of them, an active domain of them and their complementary
XX CC sequences. The invention may be useful for the production of compounds
XX CC with an antiarthritic, antiparkinsonian, neuroprotective, nootropic,
XX CC immunosuppressive, cytostatic, antipsoriatic, antiinflammatory, In
XX CC antibacterial, antiviral, antifungal or antiparasitic activity. In
XX CC addition, the disclosed sequences may be useful for gene therapy. The
XX CC polypeptides or their antibodies are useful for treating many diseases
XX CC such as arthritis, Parkinson's, Alzheimer's, autoimmune diseases, cancer,
XX CC psoriasis, inflammatory bowel disease and infections caused by bacteria,
XX CC viruses, fungi or parasites. The present sequence is that of a human
XX CC polypeptide of the invention.

XX SQ Sequence 291 AA;

Query Match 91.9%; Score 34; DB 5; Length 291;
Best Local Similarity 85.7%; Pred. NO. 1.5e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
||||:|
Db 123 ILAPPLP 129

RESULT 35
ADV58105
ID ADV58105 standard; peptide; 38 AA.
XX
AC ADV58105;
XX
DT 10-MAR-2005 (first entry)
XX
DE G protein coupled receptor peptide SEQ ID NO 5609.
XX
KW diagnosis; cancer; obesity; diabetes; asthma; inflammation; depression;
KW food; feedstuff; cosmetics; agriculture; animal breeding; GPCR.
XX
OS Unidentified.
XX
PN WO2004111636-A2.
XX
PD 23-DEC-2004.
XX
PF 17-JUN-2004; 2004WO-EP051158.
XX
PR 17-JUN-2003; 2003EP-00101775.
PR 17-JUN-2003; 2003US-0479061P.
XX
PA (VIBV-) VIB VZW.
PA (UYGE-) UNIV GENT.
XX
PI Kas K, Vandekerckhove J, Krols L;
XX WPI; 2005-057893/06.
XX
PT Identifying a peptide combo which corresponds with a family of proteins,
PT useful for diagnosing a variety of diseases, drug development or in
PT agriculture, comprises generating peptides by applying a digest on the
PT family of protein.
XX
PS Example; SEQ ID NO 5609; 265pp; English.
XX
CC The invention relates to a method of identifying a peptide combo which
CC corresponds with a family of proteins where each of the members of the
CC peptide combo is derived from a unique protein from the family. The
CC peptide combo is useful for quantifying specific known splice variants of
CC one or more particular proteins in a sample, for diagnosing complex
CC genetic diseases such as cancer, obesity, diabetes, asthma and
CC inflammation, neuropsychiatric disorders such as depression, for
CC quantifying one to several hundreds of protein disease markers
CC simultaneously leading to a more accurate diagnostic sub-classification,
CC for determining the extent of protein modification in a particular sample
CC of proteins, for tissue-typing analysis, for prenatal testing to detect
CC the presence of a congenital disease or for quantitating protein levels
CC diagnostic of a chromosomal abnormality, for diagnosing immune diseases
CC or neurological diseases, as biomarkers, preclinical drug development,
CC development of improved animal models, biomarkers related with
CC toxicology, clinical drug development, guidance marketed drugs,
CC prognostic or diagnostic disease markers, drug target validation and
CC selection, monitoring protein splicing, drug lead profiling, pathway
CC analysis, answering basic disease biology questions, and in the fields of
CC food and feed, cosmetics, agriculture and animal breeding. The present
CC sequence represents a peptide from a G-protein coupled receptor peptide
CC combo.
XX
SQ Sequence 38 AA;

Query Match 89.2%; Score 33; DB 9; Length 38;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAPPVP 7
| | | | |
Db 22 LAPPVP 27

RESULT 37

RESULT 36
ADV59047
ID ADV59047 standard; peptide; 39 AA.
XX
AC ADV59047;
XX
DT 10-MAR-2005 (first entry)
XX
DE G protein coupled receptor peptide SEQ ID NO 6552.
XX
KW diagnosis; cancer; obesity; diabetes; asthma; inflammation; depression;
KW food; feedstuff; cosmetics; agriculture; animal breeding; GPCR.
XX
OS Unidentified.
XX
PN WO2004111636-A2.
XX
PD 23-DEC-2004.
XX
PF 17-JUN-2004; 2004WO-EP051158.
XX
PR 17-JUN-2003; 2003EP-00101775.
PR 17-JUN-2003; 2003US-0479061P.
XX
PA (VIBV-) VIB VZW.
PA (UYGE-) UNIV GENT.
XX
PI Kas K, Vandekerckhove J, Krols L;
XX WPI; 2005-057893/06.
XX
PT Identifying a peptide combo which corresponds with a family of proteins,
PT useful for diagnosing a variety of diseases, drug development or in
PT agriculture, comprises generating peptides by applying a digest on the
PT family of protein.
XX
PS Example; SEQ ID NO 6552; 265pp; English.
XX
CC The invention relates to a method of identifying a peptide combo which
CC corresponds with a family of proteins where each of the members of the
CC peptide combo is derived from a unique protein from the family. The
CC peptide combo is useful for quantifying specific known splice variants of
CC one or more particular proteins in a sample, for diagnosing complex
CC genetic diseases such as cancer, obesity, diabetes, asthma and
CC inflammation, neuropsychiatric disorders such as depression, for
CC quantifying one to several hundreds of protein disease markers
CC simultaneously leading to a more accurate diagnostic sub-classification,
CC for determining the extent of protein modification in a particular sample
CC of proteins, for tissue-typing analysis, for prenatal testing to detect
CC the presence of a congenital disease or for quantitating protein levels
CC diagnostic of a chromosomal abnormality, for diagnosing immune diseases
CC or neurological diseases, as biomarkers, preclinical drug development,
CC development of improved animal models, biomarkers related with
CC toxicology, clinical drug development, guidance marketed drugs,
CC prognostic or diagnostic disease markers, drug target validation and
CC selection, monitoring protein splicing, drug lead profiling, pathway
CC analysis, answering basic disease biology questions, and in the fields of
CC food and feed, cosmetics, agriculture and animal breeding. The present
CC sequence represents a peptide from a G-protein coupled receptor peptide
CC combo.
XX
SQ Sequence 39 AA;

Query Match 89.2%; Score 33; DB 9; Length 39;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAPPVP 7
| | | | |
Db 22 LAPPVP 27

ADC94833
ID ADC94833 standard; protein; 67 AA.
XX
AC ADC94833;
XX
DT 01-JAN-2004 (first entry)
XX
DE E. faecium protein sequence SEQ ID 4460.
XX
KW Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;
KW abdominal-pelvic infection.
XX
OS Enterococcus faecium.
XX
PN US6583275-B1.
XX
PD 24-JUN-2003.
XX
PF 30-JUN-1998; 98US-00107532.
XX
PR 02-JUL-1997; 97US-0051571P.
PR 14-MAY-1998; 98US-0085598P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Doucette-Stamm LA, Bush D;
XX
DR WPI; 2003-799836/75.
DR N-PSDB; ADC91179.
XX
PT New isolated nucleic acid derived from Enterococcus faecium encoding an
PT Enterococcus faecium polypeptide useful for detection, prevention and
PT treatment of a pathological condition resulting from a bacterial
PT infection.
XX
PS Example 1; SEQ ID NO 4460; 243pp; English.
XX
CC The invention relates to an isolated nucleic acid derived from
CC Enterococcus faecium encoding an Enterococcus faecium polypeptide having
CC one of 10 fully defined sequences given in the (or comprising 40
CC sequential nucleotides chosen from any of the nucleic acids, its
CC complement or sequences hybridising to it). Also included are a
CC recombinant vector comprising the nucleic acid operably linked to
CC transcription regulatory element, a cell comprising the vector and a
CC single-stranded probe comprising the nucleic acid. The nucleic acids are
CC chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.
CC The nucleic acids is useful for diagnosing pathological conditions
CC resulting from E. faecium bacterial infection (e.g. urinary tract
CC infection, bacteraemia, endocarditis, wounds and abdominal-pelvic
CC infection) and for screening drugs such as agonists and antagonists. The
CC nucleic acid is useful for recombinant production of Candida albicans -
CC derived peptides or antisense polypeptides. Pharmaceutical compositions
CC and vaccines containing the nucleic acid are useful for preventing or
CC treating Enterococcus faecium infections. The present sequence represents
CC one if the disclosed E. faecium proteins.
XX
SQ Sequence 67 AA;

Query Match 89.2%; Score 33; DB 7; Length 67;
Best Local Similarity 85.7%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
| | | | |
Db 39 IFAPPVP 45

RESULT 38
AAY00298
ID AAY00298 standard; protein; 93 AA.
XX
AC AAY00298;
XX

DT 04-MAY-1999 (first entry)
XX
DE Human secreted protein encoded by gene 41.
XX
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS Homo sapiens.
XX
PN WO9906423-A1.
XX
PD 11-FEB-1999.
XX
PF 29-JUL-1998; 98WO-US015949.
XX
PR 30-JUL-1997; 97US-0054209P.
PR 30-JUL-1997; 97US-0054211P.
PR 30-JUL-1997; 97US-0054212P.
PR 30-JUL-1997; 97US-0054213P.
PR 30-JUL-1997; 97US-0054214P.
PR 30-JUL-1997; 97US-0054215P.
PR 30-JUL-1997; 97US-0054217P.
PR 30-JUL-1997; 97US-0054218P.
PR 30-JUL-1997; 97US-0054234P.
PR 30-JUL-1997; 97US-0054236P.
PR 18-AUG-1997; 97US-0055968P.
PR 18-AUG-1997; 97US-0055969P.
PR 18-AUG-1997; 97US-0055972P.
PR 19-AUG-1997; 97US-0056534P.
PR 19-AUG-1997; 97US-0056543P.
PR 19-AUG-1997; 97US-0056554P.
PR 19-AUG-1997; 97US-0056561P.
PR 19-AUG-1997; 97US-0056727P.
PR 19-AUG-1997; 97US-0056729P.
PR 19-AUG-1997; 97US-0056730P.
XX
(HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Feng P, Lafleur DW, Moore PA, Shi Y, Kyaw H, Li Y;
PI Zeng Z, Carter KC, Endress GA, Wei Y, Fan P, Rosen CA;
XX
DR WPI; 1999-153691/13.
DR N-PSDB; AAX20452.
XX
PT New isolated human genes and the secreted polypeptides they encode -
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders.
XX
PS Claim 11; Page 273; 312pp; English.
XX
CC This sequence represents a secreted human protein encoded by the gene
CC clone detailed in the descriptor line. The gene can be used to generate
CC fusion proteins by linking to the gene to a human immunoglobulin Fc
CC portion (e.g. AAX20403) for increasing the stability of the fused protein
CC as compared to the human protein only. The invention relates to 83 novel
CC genes and their fragments (nucleic acid sequences: AAX20412-X20499; amino
CC acid sequences AAY00258-Y00377) which are useful for preventing, treating
CC or ameliorating medical conditions e.g. by protein or gene therapy. Also,
CC pathological conditions can be diagnosed by determining the amount of the
CC new polypeptides in a sample or by determining the presence of mutations
CC in the new polynucleotides. Specific uses are described for each of the
CC 83 polynucleotides, based on which tissues they are most highly expressed
CC in (see AAX20412 for described uses)
XX
SQ Sequence 93 AA;

Query Match 89.2%; Score 33; DB 2; Length 93;

	Best Local Similarity 100.0%; Pred. No. 7.5e+02;				
	Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	2 LAPPVP 7				
Db	8 LAPPVP 13				
RESULT 39					
ADD90327					
ID	ADD90327 standard; protein; 93 AA.				
XX					
AC	ADD90327;				
XX					
DT	29-JAN-2004 (first entry)				
XX					
DE	Novel human secreted protein seq id 142.				
XX					
KW	gene therapy; cytostatic; cancer; human; secreted protein.				
XX					
OS	Homo sapiens.				
XX					
PN	US2003199683-A1.				
XX					
PD	23-OCT-2003.				
XX					
PF	30-MAR-2001; 2001US-00820649.				
XX					
PR	30-JUL-1997; 97US-0054209P.				
PR	30-JUL-1997; 97US-0054211P.				
PR	30-JUL-1997; 97US-0054212P.				
PR	30-JUL-1997; 97US-0054213P.				
PR	30-JUL-1997; 97US-0054214P.				
PR	30-JUL-1997; 97US-0054215P.				
PR	30-JUL-1997; 97US-0054217P.				
PR	30-JUL-1997; 97US-0054218P.				
PR	30-JUL-1997; 97US-0054234P.				
PR	30-JUL-1997; 97US-0054236P.				
PR	18-AUG-1997; 97US-0055968P.				
PR	18-AUG-1997; 97US-0055969P.				
PR	18-AUG-1997; 97US-0055972P.				
PR	19-AUG-1997; 97US-0056534P.				
PR	19-AUG-1997; 97US-0056543P.				
PR	19-AUG-1997; 97US-0056554P.				
PR	19-AUG-1997; 97US-0056561P.				
PR	19-AUG-1997; 97US-0056727P.				
PR	19-AUG-1997; 97US-0056729P.				
PR	29-JUL-1998; 98WO-US015949.				
PR	26-JAN-1999; 99US-00236557.				
PR	21-SEP-2000; 2000US-00666987.				
XX					
PA	(RUBE/) RUBEN S M.				
PA	(FENG/) FENG P.				
PA	(LAF/) LAFLEUR D W.				
PA	(MOOR/) MOORE P A.				
PA	(SHIY/) SHI Y.				
PA	(KYAW/) KYAW H.				
PA	(LIYY/) LI Y.				
PA	(ZENG/) ZENG Z.				
PA	(CART/) CARTER K C.				
PA	(ENDR/) ENDRESS G A.				
PA	(WEIY/) WEI Y.				
PA	(FANP/) FAN P.				
PA	(ROSE/) ROSEN C A.				
XX					
PI	Ruben SM, Feng P, Lafleur DW, Moore PA, Shi Y, Kyaw H, Li Y;				
PI	Zeng Z, Carter KC, Endress GA, Wei Y, Fan P, Rosen CA;				
XX					
DR	WPI; 2003-852813/79.				
DR	N-PSDB; ADD90236.				
XX					
PT	New nucleic acid molecule, useful for preparing a medicament for				
PT					
PT	preventing, treating or ameliorating a medical condition e.g., cancer.				
XX					
PS	Claim 11; SEQ ID NO 142; 213pp; English.				
XX					
CC	The invention describes novel isolated human nucleic acids. The nucleic				
CC	acid is useful for preparing a medicament for preventing, treating or				
CC	ameliorating a medical condition e.g., cancer, and in gene therapy. This				
CC	is the amino acid sequence of a novel human secreted protein of the				
CC	invention.				
XX					
SQ	Sequence 93 AA;				
	Query Match 89.2%; Score 33; DB 7; Length 93;				
	Best Local Similarity 100.0%; Pred. No. 7.5e+02;				
	Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	2 LAPPVP 7				
Db	8 LAPPVP 13				
RESULT 40					
ADG90146					
ID	ADG90146 standard; protein; 93 AA.				
XX					
AC	ADG90146;				
XX					
DT	11-MAR-2004 (first entry)				
XX					
DE	Human protein from secreted protein gene 41.				
XX					
KW	Secreted protein; gene therapy; neural disorder; immune system disorders;				
KW	muscular disorder; reproductive disorder; gastrointestinal disorder;				
KW	pulmonary disorder; cardiovascular disorder; renal disorder;				
KW	proliferative disorder; cancer; systemic lupus erythematosus;				
KW	rheumatoid arthritis; multiple sclerosis; thyroiditis; anaemia;				
KW	Grave's disease; diabetes; hepatitis; asthma; allergy; nephritis;				
KW	Parkinson's disease; Alzheimer's disease; atherosclerosis;				
KW	myocardial infarction; AIDS; infection; human.				
XX					
OS	Homo sapiens.				
XX					
PN	US2003166541-A1.				
XX					
PD	04-SEP-2003.				
XX					
PF	04-JUN-2002; 2002US-00160162.				
XX					
PR	30-JUL-1997; 97US-0054209P.				
PR	30-JUL-1997; 97US-0054211P.				
PR	30-JUL-1997; 97US-0054212P.				
PR	30-JUL-1997; 97US-0054213P.				
PR	30-JUL-1997; 97US-0054214P.				
PR	30-JUL-1997; 97US-0054215P.				
PR	30-JUL-1997; 97US-0054217P.				
PR	30-JUL-1997; 97US-0054218P.				
PR	30-JUL-1997; 97US-0054234P.				
PR	30-JUL-1997; 97US-0054236P.				
PR	18-AUG-1997; 97US-0055968P.				
PR	18-AUG-1997; 97US-0055969P.				
PR	18-AUG-1997; 97US-0055972P.				
PR	19-AUG-1997; 97US-0056534P.				
PR	19-AUG-1997; 97US-0056543P.				
PR	19-AUG-1997; 97US-0056554P.				
PR	19-AUG-1997; 97US-0056561P.				
PR	19-AUG-1997; 97US-0056727P.				
PR	19-AUG-1997; 97US-0056729P.				
PR	19-AUG-1997; 97US-0056730P.				
PR	29-JUL-1998; 98WO-US015949.				
PR	26-JAN-1999; 99US-00236557.				
PR	05-JUN-2001; 2001US-0295558P.				
XX					
PA	(HUMA-) HUMAN GENOME SCI INC.				

XX
PI Ruben SM, Feng P, Lafleur DW, Moore PA, Shi Y, Kyaw H, Li Y;
PI Zeng Z, Carter KC, Endress GA, Wei Y, Fan P, Rosen CA;
XX
DR WPI; 2003-874923/81.
DR N-PSDB; ADG90055.
XX
PT Nucleic acids encoding 83 secreted polypeptides, useful for preventing,
PT diagnosing and treating disorders related to their aberrant expression
PT and activity.
XX
PS Claim 11; SEQ ID NO 142; 308pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule encoding a
CC secreted protein that is at least 95% identical to a polynucleotide
CC fragment of any of the nucleotide sequences listed in table 1A of the
CC specification, which is hybridisable to the nucleotide sequences, a
CC polynucleotide encoding a polypeptide (or a polypeptide fragment, domain
CC or epitope of any of the amino acid sequences) listed in table 1A of the
CC specification, a polynucleotide which is an (allelic) variant of the
CC nucleotide sequences listed in the specification, a polynucleotide which
CC encodes a species homologue of the above amino acid sequences, a
CC polynucleotide capable of hybridising under stringent conditions to any
CC of the above polynucleotides, where the polynucleotide does not hybridise
CC under stringent conditions to a nucleic acid molecule having a nucleotide
CC sequence of only A or T residues. Also included are a recombinant vector
CC comprising the above nucleic acid molecule, making a recombinant host
CC cell comprising the above nucleic acid molecule, an isolated polypeptide
CC comprising a sequence that is at least 95% identical to the polypeptide
CC (or its fragment, domain, epitope, secreted form, (allelic) variant or
CC homologue) encoded by the above nucleic acid molecule, an isolated
CC antibody that binds specifically to the above polypeptide, a recombinant
CC host cell produced by the above method and that expresses the above
CC polypeptide, making an isolated polypeptide, preventing, treating or
CC ameliorating a medical condition, diagnosing a pathological condition or
CC a susceptibility to a pathological condition in a subject, identifying a
CC binding partner to the above polypeptide, the gene corresponding to the
CC cDNA sequence given in the specification, and identifying an activity in
CC a biological assay. The nucleic acid molecule and polypeptide are useful
CC in diagnosing, preventing, prognosing or treating diseases or disorders
CC associated with aberrant expression and/or activity of the above
CC polypeptide, such as neural disorders, immune system disorders, muscular
CC disorders, reproductive disorders, gastrointestinal disorders, pulmonary
CC disorders, cardiovascular disorders, renal disorders, proliferative
CC disorders and/or cancers. In particular, these diseases are systemic
CC lupus erythematosus, rheumatoid arthritis, multiple sclerosis,
CC thyroiditis, anaemia, Grave's disease, diabetes, hepatitis, asthma,
CC allergies, nephritis, Parkinson's disease, Alzheimer's disease,
CC atherosclerosis, myocardial infarction, AIDS and infections. The methods
CC may be used for identifying agonists and antagonists of the
CC polynucleotide and polypeptide. The present sequence is a protein from
CC one of the 83 disclosed secreted protein genes.
XX
SQ Sequence 93 AA;
Query Match 89.2%; Score 33; DB 7; Length 93;
Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 LAPPVP 7
| | | | |
Db 8 LAPPVP 13
RESULT 41
ADY25486
ID ADY25486 standard; protein; 93 AA.
XX
AC ADY25486;
XX
DT 05-MAY-2005 (first entry)
XX
DE Novel human secreted protein SeqID142.

XX
KW cancer; cytostatic; arthritis; antiarthritic; asthma; antiasthmatic;
KW acquired immune deficiency syndrome; rheumatoid arthritis; antirheumatic;
KW inflammatory bowel disease; antiinflammatory; gastrointestinal-gen.;
KW sepsis; antibacterial; immunosuppressive; acne; antiseborrheic;
KW dermatological; psoriasis; antipsoriatic; atherosclerosis;
KW antiarteriosclerotic; cerebrovascular ischemia; cerebroprotective;
KW vasotropic; thrombosis; wound healing; vulnery; Alzheimers disease;
KW neuroprotective; nootropic; parkinsons disease; antiparkinsonian; autism;
KW obsessive-compulsive disorder; tranquilizer; graft versus host disease;
KW immune disorder; hematological disease; inflammation; infection;
KW hyperproliferative disorders; renal disease; nephrotropic;
KW cardiovascular disease; cardiovascular-gen.; respiratory disorder;
KW neurological disease; neuroprotective; endocrine disease;
KW reproductive disorders (general); gynecological.
XX
OS Homo sapiens.
XX
PN US2005037467-A1.
XX
PD 17-FEB-2005.
XX
PF 09-SEP-2004; 2004US-00936773.
XX
PR 30-JUL-1997; 97US-0054209P.
PR 30-JUL-1997; 97US-0054211P.
PR 30-JUL-1997; 97US-0054212P.
PR 30-JUL-1997; 97US-0054213P.
PR 30-JUL-1997; 97US-0054214P.
PR 30-JUL-1997; 97US-0054215P.
PR 30-JUL-1997; 97US-0054217P.
PR 30-JUL-1997; 97US-0054218P.
PR 30-JUL-1997; 97US-0054223P.
PR 30-JUL-1997; 97US-0054236P.
PR 18-AUG-1997; 97US-0055968P.
PR 18-AUG-1997; 97US-0055969P.
PR 18-AUG-1997; 97US-0055972P.
PR 19-AUG-1997; 97US-0056534P.
PR 19-AUG-1997; 97US-0056543P.
PR 19-AUG-1997; 97US-0056554P.
PR 19-AUG-1997; 97US-0056561P.
PR 19-AUG-1997; 97US-0056727P.
PR 19-AUG-1997; 97US-0056729P.
PR 19-AUG-1997; 97US-0056730P.
PR 29-JUL-1998; 98WO-US015949.
PR 26-JAN-1999; 99US-00236557.
PR 21-SEP-2000; 2000US-00666984.
PR 30-MAR-2001; 2001US-00820649.
PR 05-JUN-2001; 2001US-0295558P.
PR 04-JUN-2002; 2002US-00160162.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Feng P, Lafleur DW, Moore PA, Shi Y, Kyaw H, Li Y;
PI Zeng Z, Carter KC, Endress GA, Wei Y, Fan P, Rosen CA;
XX
DR WPI; 2005-161941/17.
DR N-PSDB; ADY25395.
XX
PT New nucleic acid molecule encoding human secreted protein, useful for
PT preventing, treating, or ameliorating immune system, blood, inflammatory,
PT infectious, cardiovascular, respiratory, neurological, endocrine, or
PT reproductive disorders.
XX
PS Claim 11; SEQ ID NO 142; 309pp; English.
XX
CC The invention comprises the amino acid and coding sequences of novel
CC human secreted proteins and peptides. The DNA and protein sequences of
CC the invention are useful for treating or ameliorating: cancer, arthritis,
CC asthma, AIDS, rheumatoid arthritis, inflammatory bowel disease, sepsis,
CC acne, psoriasis, atherosclerosis, stroke, thrombosis, wound healing,
CC Alzheimer's Disease, Parkinson's Disease, autism, obsessive compulsive
CC disorder, graft-versus-host diseases, immune system disorders, blood

CC disorders, inflammatory conditions, infectious diseases,
CC hyperproliferative disorders, renal disorders, cardiovascular disorders,
CC respiratory disorders, neurological disorders, endocrine disorders, or
CC reproductive disorders. The present sequence is that of a novel human
CC secreted protein of the invention.
XX
SQ Sequence 93 AA;

Query Match 89.2%; Score 33; DB 9; Length 93;
Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAPPVP 7
Db 8 LAPPVP 13
|||||

RESULT 42
ADFF60277
ID ADF60277 standard; protein; 95 AA.
XX
AC ADF60277;
XX
DT 12-FEB-2004 (first entry)
XX
DE Human contig polypeptide sequence SEQ ID NO:2644.
XX

KW biological activity; genetic engineering; hybridisation probe; oligomer;
KW primer; chromosome mapping; gene mapping; recombinant protein production;
KW human.
XX

OS Homo sapiens.
XX
FN WO2003080795-A2.
XX
PD 02-OCT-2003.
XX
PF 09-AUG-2002; 2002WO-US025485.
XX
PR 09-AUG-2001; 2001US-0311261P.
XX
PA (HYSE-) HYSEQ INC.

PI Tang YT, Yang Y, Wang Z, Weng G, Ma Y;
XX
DR WPI; 2003-876918/81.
XX
DR N-PSDB; ADF59825.
XX
PT New polynucleotides, useful as hybridization probes, oligomers or
PT primers, for chromosome or gene mapping, for the recombinant production
PT of proteins, and for generating antisense DNA or RNA.
XX

PS Example 3; SEQ ID NO 2644; 571pp; English.

XX
CC The present invention describes isolated polynucleotide sequences (I),
CC which encode polypeptides (II) with biological activity. Also described:
CC (1) a vector comprising (I); (2) an expression vector comprising (I); (3)
CC a host cell genetically engineered to comprise (I) which is operatively
CC associated with a regulatory sequence that modulates expression of (I) in
CC the host cell; (4) a polypeptide (II) encoded by (I); (5) a composition
CC comprising the polypeptide of (4) and a carrier; (6) an antibody directed
CC against the polypeptide of (4); (7) detecting (I) or the polypeptide of
CC (4) in a sample; (8) identifying a compound that binds to the polypeptide
CC of (4); (9) producing the polypeptide of (4); and (10) a collection of
CC polynucleotides comprising at least one of the polynucleotide sequences
CC (I). The polynucleotides (I) can be used as hybridisation probes,
CC oligomers or primers, for chromosome or gene mapping, for the recombinant
CC production of proteins, and for generating antisense DNA or RNA. The
CC present sequence represents a human contig polypeptide sequence, which is
CC used in an example from the present invention.

XX Sequence 95 AA;

PR 01-SEP-2000; 2000US-0229343P.

Query Match 89.2%; Score 33; DB 7; Length 95;
Best Local Similarity 85.7%; Pred. No. 7.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
Db 53 ILAPPTP 59
|||||

RESULT 43
ABBI7509
ID ABB17509 standard; protein; 112 AA.
XX
AC ABB17509;

DT 23-JAN-2002 (first entry)

XX Human nervous system related polypeptide SEQ ID NO 6166.

KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulneryary;
KW antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine.

XX Homo sapiens.

XX WO200159063-A2.

XX 16-AUG-2001.

PF 17-JAN-2001; 2001WO-US001334.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 20-OCT-2000; 2000US-0242221P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.

PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-541565/60.
DR N-PSDB; ABA13835.
DR
XX
PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system cancers
PT and metastases.
XX
PS Claim 11; SEQ ID NO 6166; 1701pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC (ABB14678-ABB18001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, diabetes mellitus, Crohn's disease,
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 112 AA;

Query Match 89.2%; Score 33; DB 4; Length 112;
Best Local Similarity 85.7%; Pred. No. 8.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ILAPPVP 7
| | | | |
Db 49 ILAPPAP 55

RESULT 44
ABO68453
ID ABO68453 standard; protein; 161 AA.
XX
AC ABO68453;

XX 29-JUL-2004 (first entry)
XX Pseudomonas aeruginosa polypeptide #628.
DE Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
KW Pseudomonas aeruginosa.
XX US6551795-B1.
OS 22-APR-2003.
XX 18-FEB-1999; 99US-00252991.
PD 18-FEB-1998; 98US-0074788P.
XX 27-JUL-1998; 98US-0094190P.
PR (GENO-) GENOME THERAPEUTICS CORP.
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
PI WPI; 2003-615309/58.
XX N-PSDB; ABD02024.
DR Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX Disclosure; SEQ ID NO 17199; 455pp; English.
PS The invention relates to Pseudomonas aeruginosa polypeptides and the
XX polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 161 AA;
Query Match 89.2%; Score 33; DB 7; Length 161;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 LAPPVP 7
Db |||||
23 LAPPVP 28
RESULT 45
ABG27197
ID ABG27197 standard; protein; 190 AA.
XX
AC ABG27197;
XX 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #27188.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX Homo sapiens.
OS

XX WO200175067-A2.
PN 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US008631.
PD 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
PR (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
PI WPI; 2001-639362/73.
XX N-PSDB; AAS91384.
DR New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX Claim 20; SEQ ID NO 57556; 103pp; English.
PS The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological actions in
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 190 AA;
Query Match 89.2%; Score 33; DB 4; Length 190;
Best Local Similarity 71.4%; Pred. No. 1.5e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 ILAPPVP 7
Db :|||:|
32 VLAPPLP 38
RESULT 46
ADX74877
ID ADX74877 standard; protein; 191 AA.
XX
AC ADX74877;
XX 21-APR-2005 (first entry)
DT
XX Plant full length insert polypeptide seqid 44243.
XX plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;

KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content.
XX
OS Unidentified.
XX
PN US2004034888-A1.
XX
PD 19-FEB-2004.
XX
PF 28-APR-2003; 2003US-00425114.
XX
PR 06-MAY-1999; 99US-00304517.
PR 05-NOV-2001; 2001US-00985678.
XX
PA (LIUJ/) LIU J.
PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA (TABA/) TABASKA J E.
PA (CAOY/) CAO Y.
XX
PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
DR WPI; 2004-180133/17.
XX
XX New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
PS Claim 1; SEQ ID NO 44243; 15pp; English.
XX
XX The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.seqdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This is the amino acid sequence of a plant full length insert
CC polypeptide that can be used in the recombinant DNA construct of the
CC invention.
XX
SQ Sequence 191 AA;

Query Match 89.2%; Score 33; DB 8; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAPPVP 7
Db 105 LAPPVP 110
| | | | |

RESULT 47
ADC31836
ID ADC31836 standard; protein; 198 AA.
XX
AC ADC31836;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human novel polypeptide sequence, SEQ ID NO:1918.
XX
KW Human; diagnostic; drug screening; forensics; gene mapping;

KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW ulcers; osteoporosis; autoimmune disease; cancer;
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;
KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW gene therapy.
XX
XX Homo sapiens.
OS
XX WO2003029271-A2.
PN
XX 10-APR-2003.
PD
XX 24-SEP-2002; 2002WO-US030474.
PF
XX 24-SEP-2001; 2001US-0324631P.
PR
XX (HYSE-) HYSEQ INC.
PA
XX
PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
PI Haley-Vicente D, Drmanac RT;
XX
DR WPI; 2003-371981/35.
DR N-PSDB; ADC30865.
XX
PT New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT cancer.
XX
PS Claim 20; SEQ ID NO 1918; 1185pp; English.
XX
CC The invention relates to 971 novel human cDNA sequences (ADC29919-
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody
CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of preventing, treating or
CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC -ADC33394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a specifically
CC claimed human polypeptide sequence of the invention. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 198 AA;

Query Match 89.2%; Score 33; DB 7; Length 198;
Best Local Similarity 85.7%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAPPVP 7
133 LAPPVP 138

Db

RESULT 50
ADX77177
ID ADX77177 standard; protein; 285 AA.
XX
AC ADX77177;
XX
DT 21-APR-2005 (first entry)
XX
DE Plant full length insert polypeptide seqid 46543.
XX
KW plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content.
XX
OS Unidentified.
XX
PN US2004034888-A1.
XX
PD 19-FEB-2004.
XX
PF 28-APR-2003; 2003US-00425114.
XX
PR 06-MAY-1999; 99US-00304517.
PR 05-NOV-2001; 2001US-00985678.
XX
PA (LIUJ/) LIU J.
PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA (TABAJ/) TABASKA J E.
PA (CAOY/) CAO Y.
XX
PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX
DR WPI; 2004-180133/17.
XX
PT New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
PS Claim 1; SEQ ID NO 46543; 15pp; English.
XX
CC The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.seqdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This is the amino acid sequence of a plant full length insert
CC polypeptide that can be used in the recombinant DNA construct of the
CC invention.

XX SQ Sequence 285 AA;
Query Match 89.2%; Score 33; DB 8; Length 285;
Best Local Similarity 100.0%; Pred. No. 2.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAPPVP 7
Db 67 LAPPVP 72

Search completed: April 6, 2006, 09:31:52
Job time : 114.289 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2006, 09:35:13 ; Search time 12.7105 Seconds
(without alignments)
52.989 Million cell updates/sec

Title: US-10-632-388-296

Perfect score: 37

Sequence: 1 ILAPPVP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : PIR 80:*

- 1: pir1:*
- 2: pir2:*
- 3: pir3:*
- 4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	37	100.0	1615	2	B49502	protein-tyrosine-p
2	37	100.0	1767	2	A49502	protein-tyrosine-p
3	35	94.6	266	2	F70595	hypothetical prote
4	35	94.6	441	2	T24591	hypothetical prote
5	34	91.9	189	2	D71658	3-octaprenyl-4-hyd
6	34	91.9	924	2	S75284	chemotaxis protein
7	33	89.2	113	2	AE1448	hypothetical prote
8	33	89.2	113	2	AE1729	hypothetical prote
9	33	89.2	113	2	AG1752	hypothetical prote
10	33	89.2	205	2	S41002	hypothetical prote
11	33	89.2	465	2	T35712	integral membrane
12	33	89.2	538	2	S76481	hypothetical prote
13	33	89.2	543	2	AF1927	hypothetical prote
14	33	89.2	551	2	T13994	envelope protein -
15	33	89.2	687	2	T08877	Modin - Podospora
16	33	89.2	819	2	F87708	cell division prot
17	33	89.2	981	2	A41401	mineralocorticoid
18	33	89.2	984	2	A29513	mineralocorticoid
19	33	89.2	1367	2	T21913	hypothetical prote
20	33	89.2	7962	2	I38346	elastic titin - hu
21	32	86.5	191	2	A97800	hypothetical prote
22	32	86.5	210	2	H72585	hypothetical prote
23	32	86.5	228	2	T10399	hypothetical prote
24	32	86.5	228	2	D30857	hypothetical prote
25	32	86.5	481	2	E86433	protein T17H7.3 [i
26	32	86.5	581	2	T24393	hypothetical prote
27	32	86.5	642	2	C86152	T7123.2 protein -
28	32	86.5	720	2	A96807	hypothetical prote
29	32	86.5	753	2	T01619	hypothetical prote

30	86.5	1069	2	T00377	KIAA0642 protein -
31	86.5	1610	2	D89451	protein T04G9.1 [i
32	86.5	1626	2	A75613	hypothetical prote
33	86.5	1630	2	T00390	KIAA0614 protein -
34	86.5	1791	2	T24089	hypothetical prote
35	86.5	3530	2	A59266	unconventional myo
36	83.8	162	2	H87656	3-octaprenyl-4-hyd
37	83.8	198	2	H89164	protein T06E4.8 [i
38	83.8	199	2	T24595	hypothetical prote
39	83.8	205	2	T51884	hypothetical prote
40	83.8	210	2	G95858	probable 3-octapre
41	83.8	212	2	T24594	hypothetical prote
42	83.8	213	2	G89164	protein T06E4.9 [i
43	83.8	237	2	F95271	hypothetical prote
44	83.8	239	2	T42539	hypothetical prote
45	83.8	243	1	B42768	homeotic protein g
46	83.8	243	2	I51424	homeotic protein g
47	83.8	326	2	AF2138	hypothetical prote
48	83.8	332	2	T51269	hypothetical prote
49	83.8	347	2	A96835	gibberellin 3 beta
50	83.8	347	2	T51691	gibberellin 3 beta
51	83.8	384	2	H88924	protein F33E11.2 [
52	83.8	402	2	S77057	hypothetical prote
53	83.8	410	1	IXBE10	alpha trans-induci
54	83.8	474	2	I50830	Ig mu chain - Lepi
55	83.8	486	2	A41537	DNA-binding protei
56	83.8	565	2	AC3637	3-hydroxybutyryl-C
57	83.8	619	2	T28998	hypothetical prote
58	83.8	674	2	A97490	hypothetical prote
59	83.8	877	1	A25962	mitosis inhibitor
60	83.8	958	2	T13593	hypothetical prote
61	83.8	1027	2	T46481	hypothetical prote
62	83.8	1230	2	T18256	probable serine/th
63	83.8	1230	2	T18259	serine/threonine p
64	83.8	1306	2	T13592	hypothetical prote
65	83.8	1530	1	S01393	DNA-directed RNA p
66	83.8	1547	2	JQ0096	hypothetical 176K
67	83.8	1611	2	T38236	hypothetical prote
68	83.8	2606	2	T03159	large tegument pro
69	81.1	111	2	T49579	hypothetical prote
70	81.1	115	2	T50390	hypothetical prote
71	81.1	136	2	C72524	hypothetical prote
72	81.1	156	2	T43957	hypothetical prote
73	81.1	163	2	A70847	hypothetical prote
74	81.1	177	2	JC4839	delta Spi-B transc
75	81.1	179	2	A69104	hypothetical prote
76	81.1	194	2	E87504	hypothetical prote
77	81.1	197	2	E95411	probable decarboxy
78	81.1	203	2	AG3405	hypothetical exonu
79	81.1	212	2	H87472	hypothetical prote
80	81.1	231	2	T27803	hypothetical prote
81	81.1	250	2	D75317	hypothetical prote
82	81.1	262	2	S25655	Spi-B protein - hu
83	81.1	271	2	G83554	hypothetical prote
84	81.1	275	2	T24608	hypothetical prote
85	81.1	279	2	G75498	aminoglycoside N3-
86	81.1	281	2	T24658	hypothetical prote
87	81.1	296	2	G86446	unknown protein [i
88	81.1	304	2	T04053	nodulin-26 homolog
89	81.1	363	2	AH1880	hypothetical prote
90	81.1	365	2	T04718	hypothetical prote
91	81.1	367	2	I39172	cyclin-dependent k
92	81.1	370	2	AB2490	hypothetical prote
93	81.1	377	2	T52606	squamosa promoter
94	81.1	395	2	AE2349	hypothetical prote
95	81.1	402	2	JE0282	cell division prot
96	81.1	431	2	AF2635	exopolysaccharide
97	81.1	435	2	E97417	exoQ-like protein
98	81.1	458	2	F86433	protein T17H7.5 [i
99	81.1	469	2	C87629	major facilitator
100	81.1	476	2	JC7304	fatty acid hydrope
101	81.1	477	2	A47236	zinc-finger protei
102	81.1	480	2	S74228	fatty acid hydrope

103 30 81.1 494 2 A42170 zinc finger protei
104 30 81.1 497 2 JC5076 myc-associated zin
105 30 81.1 500 2 T14826 transcription init
106 30 81.1 509 2 T02864 probable Zn finger
107 30 81.1 510 2 A45338 connexin-56 - chic
108 30 81.1 541 1 S45428 PGT112 protein - y
109 30 81.1 542 2 T23869 hypothetical prote
110 30 81.1 567 2 A86262 hypothetical prote
111 30 81.1 614 2 S58306 WD-40 repeat regul
112 30 81.1 624 1 RDCS7 sulfite reductase
113 30 81.1 634 2 T00359 hypothetical prote
114 30 81.1 641 2 T30076 hypothetical prote
115 30 81.1 643 2 E70682 probable gamma-glu
116 30 81.1 650 2 A11974 ferredoxin-sulfite
117 30 81.1 708 2 T00064 hypothetical prote
118 30 81.1 722 2 G75563 phosphate acetyltr
119 30 81.1 726 2 T35865 probable cyclase -
120 30 81.1 743 2 T13673 hypothetical prote
121 30 81.1 758 2 T15577 phosphatidylglycer
122 30 81.1 763 2 AD1070 squamosa promoter
123 30 81.1 801 2 T52605 cadherin-15 precur
124 30 81.1 814 2 G02878 verprolin - yeast
125 30 81.1 817 2 S51342 ribonuclease, Rne/
126 30 81.1 898 2 H87481 hypothetical prote
127 30 81.1 1008 2 T04462 major acidic nucle
128 30 81.1 1023 2 JC4013 formin related pro
129 30 81.1 1064 2 T13963 hypothetical prote
130 30 81.1 1069 2 D85383 hypothetical prote
131 30 81.1 1072 2 T18802 cobN protein homol
132 30 81.1 1085 2 T03531 ecdysone-induced p
133 30 81.1 1237 2 A34598 conserved hypothet
134 30 81.1 1245 2 H83574 probable ubiquitou
135 30 81.1 1401 2 T02255 poly(A)-specific r
136 30 81.1 1480 2 S48440 probable ppsB prot
137 30 81.1 1538 2 E70874 protein-tyrosine-p
138 30 81.1 1557 2 D41214 protein-tyrosine-p
139 30 81.1 1630 2 C41214 DNA-directed RNA p
140 30 81.1 1781 1 A34374 genome polypotein
141 30 81.1 3010 1 GNWVTC genome polypotein
142 30 81.1 3010 1 S18030 genome polypotein
143 29 78.4 17 2 A49237 45/47K antigen - M
144 29 78.4 29 2 C61233 conceptus protein
145 29 78.4 61 2 AD3152 hypothetical prote
146 29 78.4 67 2 G98135 hypothetical prote
147 29 78.4 92 2 S52277 hypothetical prote
148 29 78.4 108 2 G72474 hypothetical prote
149 29 78.4 115 2 T30042 hypothetical prote
150 29 78.4 122 2 T32932 hypothetical prote
151 29 78.4 124 2 S25113 insulin-like growt
152 29 78.4 127 2 PC2265 cytochrome P450 pr
153 29 78.4 127 2 B72732 hypothetical prote
154 29 78.4 128 2 T32947 hypothetical prote
155 29 78.4 131 2 A84513 hypothetical prote
156 29 78.4 133 2 E70795 hypothetical prote
157 29 78.4 136 2 G75482 hypothetical prote
158 29 78.4 138 2 E95917 hypothetical proli
159 29 78.4 152 2 AH3007 conserved hypothet
160 29 78.4 152 2 E98276 hypothetical prote
161 29 78.4 162 2 E87094 probable phosphope
162 29 78.4 162 2 H70865 probable isomerase
163 29 78.4 163 2 F86438 protein T19E23.3 [
164 29 78.4 164 2 E75293 hypothetical prote
165 29 78.4 172 2 A70864 hypothetical prote
166 29 78.4 185 2 D75602 hypothetical prote
167 29 78.4 189 1 XMECPD 3-octaprenyl-4-hyd
168 29 78.4 189 2 AB0801 probable decarboxy
169 29 78.4 189 2 C91028 3-octaprenyl-4-hyd
170 29 78.4 189 2 D85872 3-octaprenyl-4-hyd
171 29 78.4 189 2 T05383 hypothetical prote
172 29 78.4 190 2 T35570 hypothetical prote
173 29 78.4 196 2 D98190 decarboxylase [imp
174 29 78.4 196 2 AE3096 phenylacrylic acid
175 29 78.4 196 2 AC3629 3-octaprenyl-4-hyd

176 29 78.4 197 1 JH0309 tumor necrosis fac
177 29 78.4 197 2 S22010 pyrimidine synthes
178 29 78.4 197 2 T19892 hypothetical prote
179 29 78.4 204 2 T28010 hypothetical prote
180 29 78.4 215 2 T35493 hypothetical prote
181 29 78.4 216 2 T02658 probable germin pr
182 29 78.4 218 2 C86337 hypothetical prote
183 29 78.4 218 2 T26826 hypothetical prote
184 29 78.4 224 2 C48652 transfer protein s
185 29 78.4 242 2 S62017 phenylacrylic acid
186 29 78.4 243 2 D84792 probable protein t
187 29 78.4 245 2 T20831 hypothetical prote
188 29 78.4 247 2 T26493 hypothetical prote
189 29 78.4 248 2 AH2339 hypothetical prote
190 29 78.4 250 2 A54308 pg-1 protein - hum
191 29 78.4 255 2 T41451 very hypothetical
192 29 78.4 259 2 T22596 hypothetical prote
193 29 78.4 263 1 S23009 insulin-like growt
194 29 78.4 272 1 A36082 insulin-like growt
195 29 78.4 272 2 I48600 insulin-like growt
196 29 78.4 274 2 T04619 hypothetical prote
197 29 78.4 274 2 T47730 hypothetical prote
198 29 78.4 278 2 A97471 hypothetical prote
199 29 78.4 278 2 AF2689 conserved hypothet
200 29 78.4 283 2 D83948 hypothetical prote
201 29 78.4 297 2 T14281 p40-like ribosomal
202 29 78.4 304 1 XMECBD acetyl-CoA carboxy
203 29 78.4 304 2 A85873 acetyl-CoA carboxy
204 29 78.4 304 2 H91028 peptidase M20/M25/
205 29 78.4 309 2 E87559 eaA protein - phag
206 29 78.4 317 2 S35282 hrpW protein - Pse
207 29 78.4 323 2 S62085 probable modD prot
208 29 78.4 325 2 D70666 probable membrane
209 29 78.4 335 2 G64817 hypothetical prote
210 29 78.4 335 2 G85589 hypothetical prote
211 29 78.4 335 2 E90739 hypothetical prote
212 29 78.4 336 2 E47301 Vir10 homolog - Bo
213 29 78.4 336 2 T30459 hypothetical prote
214 29 78.4 345 2 AD3537 oligopeptide trans
215 29 78.4 347 2 T13891 NADH2 dehydrogenas
216 29 78.4 348 2 C69399 N5,N10-methylenete
217 29 78.4 350 1 A45663 thymidine kinase (
218 29 78.4 353 1 C54258 transcription fact
219 29 78.4 354 1 S35090 transcription fact
220 29 78.4 358 2 T13017 hypothetical prote
221 29 78.4 363 2 T16755 hypothetical prote
222 29 78.4 365 2 S56792 hypothetical prote
223 29 78.4 370 2 C83120 probable RND efflu
224 29 78.4 371 2 T49696 related to peroxis
225 29 78.4 371 2 B39625 T-cell receptor al
226 29 78.4 372 2 S69574 hypothetical prote
227 29 78.4 373 1 VGBETC glycoprotein A - t
228 29 78.4 373 2 T16529 hypothetical prote
229 29 78.4 374 2 S48639 fructose-bisphosph
230 29 78.4 378 2 T12621 Dc3 promoter-bindi
231 29 78.4 386 2 S40296 glycine amidinotra
232 29 78.4 386 2 F72773 probable molybdopt
233 29 78.4 397 2 A39565 lymphoid enhancer-
234 29 78.4 398 2 A40966 porphobilinogen sy
235 29 78.4 399 2 A39625 T-cell receptor al
236 29 78.4 402 2 D70908 probable lprL prot
237 29 78.4 407 2 C70816 hypothetical prote
238 29 78.4 409 2 T33019 hypothetical prote
239 29 78.4 411 2 A70509 hypothetical prote
240 29 78.4 412 2 T26056 hypothetical prote
241 29 78.4 413 2 AH2743 conserved hypothet
242 29 78.4 414 2 I48975 apolipoprotein A-I
243 29 78.4 414 2 A37133 apolipoprotein A-I
244 29 78.4 415 2 T34156 hypothetical prote
245 29 78.4 415 2 T26057 hypothetical prote
246 29 78.4 416 2 G97524 hypothetical prote
247 29 78.4 420 2 B75333 twitching mobility
248 29 78.4 422 2 G70984 probable papA5 pro

249 29 78.4 423 2 S73020 hypothetical prote
250 29 78.4 423 2 T18784 hypothetical prote
251 29 78.4 424 2 A50000 porphobilinogen sy
252 29 78.4 430 2 S58169 porphobilinogen sy
253 29 78.4 430 2 D96719 hypothetical prote
254 29 78.4 447 1 A70773 diaminopimelate de
255 29 78.4 458 2 T06313 hypothetical prote
256 29 78.4 464 2 T36256 probable membrane
257 29 78.4 465 2 S68987 transcription acti
258 29 78.4 467 2 A59268 nodulation competi
259 29 78.4 468 2 S26741 T-cell glycoprotei
260 29 78.4 472 2 S39507 glucuronosyl trans
261 29 78.4 478 1 S03826 UMP synthase - gli
262 29 78.4 482 2 T36045 hypothetical prote
263 29 78.4 486 2 AG2199 hypothetical prote
264 29 78.4 488 2 F96724 hypothetical prote
265 29 78.4 489 1 VGBETA glycoprotein A - t
266 29 78.4 494 2 T39763 mitochondrial proc
267 29 78.4 495 2 A40465 alanine transamina
268 29 78.4 495 2 A39900 alanine transamina
269 29 78.4 501 1 A60005 glycoprotein A pre
270 29 78.4 501 1 VGBEMA glycoprotein A pre
271 29 78.4 501 1 VGBEMB glycoprotein A pre
272 29 78.4 502 2 T00483 hypothetical prote
273 29 78.4 503 2 S36549 l1 protein - human
274 29 78.4 505 1 VGBEMH glycoprotein gp57-
275 29 78.4 505 2 D69382 conserved hypothet
276 29 78.4 509 2 T29291 hypothetical prote
277 29 78.4 514 2 C49507 potassium channel
278 29 78.4 515 2 B83477 hypothetical prote
279 29 78.4 517 2 A48250 U2AF-homologous pr
280 29 78.4 520 2 T26201 hypothetical prote
281 29 78.4 523 1 A60408 glycoprotein gp57-
282 29 78.4 525 2 T20662 hypothetical prote
283 29 78.4 541 2 E87414 transglycosylase,
284 29 78.4 548 2 T23270 acetyl choline rec
285 29 78.4 548 2 S59133 Ets2 repressor fac
286 29 78.4 553 1 G1BPSV gene 1 protein - s
287 29 78.4 562 2 AD0304 probable hemolysin
288 29 78.4 562 2 AD0074 probable hemolysin
289 29 78.4 574 2 T43556 Wiskott-Aldrich sy
290 29 78.4 574 2 T38819 wiskott-aldrich sy
291 29 78.4 575 2 S59740 NRD1 protein - yea
292 29 78.4 579 2 E75275 hypothetical prote
293 29 78.4 590 2 S16411 terminase ATPase c
294 29 78.4 602 2 A49507 potassium channel
295 29 78.4 605 2 T27397 hypothetical prote
296 29 78.4 607 2 T34040 hypothetical prote
297 29 78.4 609 2 T01892 hypothetical prote
298 29 78.4 611 2 T21747 hypothetical prote
299 29 78.4 618 2 B87564 conserved hypothet
300 29 78.4 621 2 G83474 hypothetical prote
301 29 78.4 623 2 G70644 probable sppA prot
302 29 78.4 625 2 S69707 hypothetical prote
303 29 78.4 628 2 A84087 hypothetical prote
304 29 78.4 636 2 T38010 myc far upstream e
305 29 78.4 644 2 A53184 hypothetical prote
306 29 78.4 662 2 T04856 hypothetical prote
307 29 78.4 678 2 A75580 hypothetical prote
308 29 78.4 684 2 T25603 hypothetical prote
309 29 78.4 685 2 S64158 hypothetical prote
310 29 78.4 693 2 A41617 dopamine transport
311 29 78.4 717 1 S29923 transforming prote
312 29 78.4 723 2 B55926 DNA binding protei
313 29 78.4 731 2 T19721 hypothetical prote
314 29 78.4 732 1 A35655 peptidyl-dipeptida
315 29 78.4 732 1 S05238 peptidyl-dipeptida
316 29 78.4 737 2 S28030 DNA-binding protei
317 29 78.4 743 2 C56695 transducin-like en
318 29 78.4 745 2 S13586 triacylglycerol li
319 29 78.4 750 2 D86245 hypothetical prote
320 29 78.4 750 2 AH3158 hypothetical prote
321 29 78.4 751 2 T34490 hypothetical prote

322 29 78.4 762 2 T14815 hypothetical prote
323 29 78.4 766 2 S69657 hypothetical prote
324 29 78.4 771 2 S56791 hypothetical prote
325 29 78.4 798 2 F96714 probable protease
326 29 78.4 802 2 H59434 oligophrenin 1, Rh
327 29 78.4 810 2 C70791 probable ponA' pro
328 29 78.4 832 2 F98128 insertion element
329 29 78.4 844 2 S61104 BRO1 protein - yea
330 29 78.4 848 2 I55498 testicular dynamin
331 29 78.4 851 2 B40671 dynamin, internal
332 29 78.4 851 2 S11508 D100 protein - rat
333 29 78.4 864 2 A40671 dynamin, internal
334 29 78.4 868 2 A30817 serendipity (sry h
335 29 78.4 868 2 AF3204 autotransporter pr
336 29 78.4 877 2 S49197 envelope protein p
337 29 78.4 896 2 S36326 clathrin assembly
338 29 78.4 901 2 A44825 phosphoprotein, sy
339 29 78.4 903 2 T00074 hypothetical prote
340 29 78.4 915 2 S36327 clathrin assembly
341 29 78.4 939 2 C70876 hypothetical prote
342 29 78.4 955 2 F84914 hypothetical prote
343 29 78.4 963 2 T19140 hypothetical prote
344 29 78.4 992 2 G70950 hypothetical prote
345 29 78.4 1011 2 T07712 probable ABC-type
346 29 78.4 1013 2 T33470 hypothetical prote
347 29 78.4 1018 2 T22318 hypothetical prote
348 29 78.4 1023 2 B59430 KIAA0189 protein l
349 29 78.4 1048 2 T31425 C-terminal domain-
350 29 78.4 1087 2 T49496 hypothetical prote
351 29 78.4 1114 2 T49517 p63 related protei
352 29 78.4 1140 2 D88690 protein F41H10.3 l
353 29 78.4 1175 2 T20346 pyruvate carboxyla
354 29 78.4 1186 2 C64588 cag pathogenicity
355 29 78.4 1193 2 JC2489 peptidyl-dipeptida
356 29 78.4 1213 2 A41724 limb deformity (ld
357 29 78.4 1218 2 S71376 glutamate receptor
358 29 78.4 1223 2 T10365 helicase - Orgyia
359 29 78.4 1232 2 S40766 hypothetical prote
360 29 78.4 1236 2 E70977 hypothetical prote
361 29 78.4 1297 2 S25714 son-of-sevenless-2
362 29 78.4 1306 1 A31759 peptidyl-dipeptida
363 29 78.4 1307 2 G96711 unknown protein, 9
364 29 78.4 1307 2 T17453 ERG-associated pro
365 29 78.4 1312 1 A34171 peptidyl-dipeptida
366 29 78.4 1313 1 JC2038 peptidyl-dipeptida
367 29 78.4 1360 2 T34302 cell polarity prot
368 29 78.4 1367 2 T33819 hypothetical prote
369 29 78.4 1413 2 T26467 hypothetical prote
370 29 78.4 1417 2 A57570 Bloom's syndrome r
371 29 78.4 1457 2 T14577 protein kinase Yak
372 29 78.4 1469 2 T09219 basal transcriptio
373 29 78.4 1490 2 S72351 nonstructural poly
374 29 78.4 1533 2 A46221 abdominal segment
375 29 78.4 1634 2 T26517 hypothetical prote
376 29 78.4 1872 2 T24683 hypothetical prote
377 29 78.4 2022 2 T48818 glucan 1,4-alpha-g
378 29 78.4 2150 2 T08165 RNA1 polyprotein -
379 29 78.4 2225 1 A23443 pyrimidine synthes
380 29 78.4 2240 2 T37057 probable multi-dom
381 29 78.4 2297 2 T34918 polyketide synthas
382 29 78.4 2344 2 T41590 probable sensor-li
383 29 78.4 2493 2 S72349 nonstructural poly
384 29 78.4 2493 2 S26372 nonstructural poly
385 29 78.4 2970 2 T08839 polyprotein - marm
386 29 78.4 3005 2 T08841 polyprotein - dour
387 29 78.4 3010 1 GNWVTW genome polyprotein
388 29 78.4 4861 2 S71752 giant protein p619
389 29 78.4 13055 2 T16580 hypothetical prote
390 28 75.7 12 2 PA0098 ribosomal protein
391 28 75.7 62 2 S71752 giant protein p619
392 28 75.7 62 2 T36447 hypothetical prote
393 28 75.7 63 2 F83293 hypothetical prote
394 28 75.7 90 2 D24356 hypothetical prote

395 28 75.7 92 2 S49385 movement protein V
396 28 75.7 101 2 H75370 stage V sporulatio
397 28 75.7 105 2 H72736 hypothetical prote
398 28 75.7 107 1 WMBEL2 latency-related pr
399 28 75.7 117 2 F70881 hypothetical prote
400 28 75.7 118 2 T26804 hypothetical prote
401 28 75.7 121 2 S65808 transcription acti
402 28 75.7 124 2 T26799 hypothetical prote
403 28 75.7 127 2 PC2266 cytochrome P450 pr
404 28 75.7 138 2 B72469 hypothetical prote
405 28 75.7 140 2 T35301 hypothetical prote
406 28 75.7 141 2 T19578 hypothetical prote
407 28 75.7 151 2 D87663 hypothetical prote
408 28 75.7 153 2 T52498 hypothetical prote
409 28 75.7 158 2 S46257 mitochondrial inne
410 28 75.7 160 2 S09823 hypothetical prote
411 28 75.7 164 2 T26608 hypothetical prote
412 28 75.7 174 2 B72239 conserved hypothet
413 28 75.7 184 2 F64312 phenylacrylic acid
414 28 75.7 185 2 T00519 proline-rich prote
415 28 75.7 186 2 T45838 hypothetical prote
416 28 75.7 190 2 T26609 hypothetical prote
417 28 75.7 200 2 B95349 Probable decarboxy
418 28 75.7 202 2 JN0579 DNA-binding protei
419 28 75.7 204 2 T01376 hypothetical prote
420 28 75.7 204 2 T31290 3-octaprenyl-4-hyd
421 28 75.7 207 2 C86147 T1N6.2 protein - A
422 28 75.7 226 2 S77174 nitrate/nitrite re
423 28 75.7 227 2 D85018 probable hypoersen
424 28 75.7 229 2 S57957 thrombospondin 1 -
425 28 75.7 244 2 T37686 hypothetical prote
426 28 75.7 245 2 S63455 hypothetical prote
427 28 75.7 252 2 I46859 MHC RLA - rabbit (
428 28 75.7 256 2 H96524 protein T1N15.11 [
429 28 75.7 257 2 B84860 hypothetical prote
430 28 75.7 259 2 T36003 hypothetical prote
431 28 75.7 260 2 G87349 conserved hypothet
432 28 75.7 260 2 A38114 resolvase rsd - Sa
433 28 75.7 262 2 T01562 hypothetical prote
434 28 75.7 262 2 T48954 hypothetical prote
435 28 75.7 263 2 D71342 hypothetical prote
436 28 75.7 268 2 A24716 D protein - Escher
437 28 75.7 268 2 T00239 resolvase - Escher
438 28 75.7 269 2 T26957 hypothetical prote
439 28 75.7 271 2 T40982 hypothetical prote
440 28 75.7 275 2 G75606 probable glycohydr
441 28 75.7 281 2 D70845 hypothetical prote
442 28 75.7 283 2 E88597 protein Y47D3B.6 [
443 28 75.7 297 2 S23737 proline-rich prote
444 28 75.7 299 2 S60521 envelope polyprote
445 28 75.7 299 2 S60552 envelope polyprote
446 28 75.7 299 2 S60551 envelope polyprote
447 28 75.7 299 2 S60523 envelope polyprote
448 28 75.7 299 2 S60554 envelope polyprote
449 28 75.7 299 2 S60553 envelope polyprote
450 28 75.7 300 1 S57749 SURF1 protein - hu
451 28 75.7 300 2 AG0983 dipeptide transpor
452 28 75.7 300 2 S60522 envelope polyprote
453 28 75.7 302 2 T45937 hypothetical prote
454 28 75.7 303 2 S60549 envelope polyprote
455 28 75.7 303 2 S60550 envelope polyprote
456 28 75.7 304 2 F75278 conserved hypothet
457 28 75.7 305 2 T19710 hypothetical prote
458 28 75.7 310 2 T33611 hypothetical prote
459 28 75.7 312 2 A56911 TRADD protein - hu
460 28 75.7 315 2 A83496 conserved hypothet
461 28 75.7 316 2 D75588 transcription regu
462 28 75.7 316 2 T34023 hypothetical prote
463 28 75.7 321 2 T03431 probable transport
464 28 75.7 321 2 AB3245 hypothetical prote
465 28 75.7 334 2 E83122 probable iron/asco
466 28 75.7 335 2 S09275 Ig alpha chain C r
467 28 75.7 336 2 A83166 conserved hypothet

468 28 75.7 336 2 AG0307 conserved hypothet
469 28 75.7 339 2 E83419 probable permease
470 28 75.7 340 1 QQBES6 UL20 protein precu
471 28 75.7 343 2 S09272 Ig alpha chain C r
472 28 75.7 345 2 B75274 conserved hypothet
473 28 75.7 348 2 D88088 protein B0454.1 li
474 28 75.7 350 2 B82713 lipase modulator X
475 28 75.7 356 2 T51944 pathogenicity MAP
476 28 75.7 357 2 S09265 Ig alpha chain C r
477 28 75.7 359 1 K1BBBT thymidine kinase (
478 28 75.7 361 2 T51943 mitogen-activated
479 28 75.7 361 2 JQ2379 homeobox 1 protein
480 28 75.7 362 2 H82178 conserved hypothet
481 28 75.7 369 2 E81722 hypothetical prote
482 28 75.7 374 2 G84596 hypothetical prote
483 28 75.7 383 2 H75263 probable zinc prot
484 28 75.7 396 2 S22909 interleukin-3 rece
485 28 75.7 403 2 S52796 prpL2 protein - hu
486 28 75.7 407 2 T00693 hypothetical prote
487 28 75.7 418 2 A37344 acrosin (EC 3.4.21
488 28 75.7 420 2 A55283 acrosin (EC 3.4.21
489 28 75.7 426 2 JC7773 IL-13Ralpha 1 prot
490 28 75.7 428 2 AH3536 dihydroliipoamide S
491 28 75.7 433 2 T36122 probable integral
492 28 75.7 433 2 A10536 probable membrane
493 28 75.7 434 2 S62168 hypothetical prote
494 28 75.7 436 2 JX0172 acrosin (EC 3.4.21
495 28 75.7 436 2 JQ1459 Bt1 protein precur
496 28 75.7 439 2 S58327 cobalt accumulatio
497 28 75.7 448 2 JN0118 glucan 1,3-beta-gl
498 28 75.7 449 2 T02625 hypothetical prote
499 28 75.7 452 2 JC4100 hydroxyindole O-me
500 28 75.7 452 2 S46798 hypothetical prote
501 28 75.7 469 2 AE2188 hypothetical prote
502 28 75.7 472 2 T03169 probable glycoprot
503 28 75.7 473 1 RGBYM3 regulatory protein
504 28 75.7 478 2 I53960 PRR2 alpha - human
505 28 75.7 489 2 B33416 nuclear factor I -
506 28 75.7 491 2 T07598 proline-rich prote
507 28 75.7 494 2 AC1742 protein involved i
508 28 75.7 494 2 AE1372 proteins involved
509 28 75.7 497 2 T27012 hypothetical prote
510 28 75.7 498 2 F89861 Na+/H+-antiporter
511 28 75.7 499 2 S01038 transcription fact
512 28 75.7 504 2 G02474 interferon regulat
513 28 75.7 506 2 S37156 NADPH-ferrihemopro
514 28 75.7 506 2 A33416 nuclear factor I -
515 28 75.7 510 2 S45571 nuclear factor I-C
516 28 75.7 533 1 TVFVFP protein-tyrosine k
517 28 75.7 538 2 I68093 PRR2 delta - human
518 28 75.7 538 2 T22654 hypothetical prote
519 28 75.7 547 2 C96828 unknown protein F1
520 28 75.7 547 2 S77410 translation releas
521 28 75.7 554 2 S75016 phosphoglycerate d
522 28 75.7 554 2 F86244 hypothetical prote
523 28 75.7 557 2 T41495 hypothetical prote
524 28 75.7 560 2 F84341 hypothetical prote
525 28 75.7 572 2 H96685 probable AMP-bindi
526 28 75.7 574 2 T05964 probable low-affin
527 28 75.7 575 2 B71525 hypothetical prote
528 28 75.7 576 2 T18067 hypothetical prote
529 28 75.7 582 2 S57722 sli-1 protein - Ca
530 28 75.7 585 2 T19814 two-component resp
531 28 75.7 588 2 AE1926 probable polypepti
532 28 75.7 589 2 T42244 probable membrane-
533 28 75.7 598 2 E72733 hypothetical prote
534 28 75.7 600 2 T24447 hypothetical prote
535 28 75.7 621 2 T24090 hypothetical prote
536 28 75.7 631 2 T05016 somatotropin recep
537 28 75.7 634 2 S33339 hypothetical prote
538 28 75.7 635 2 F75477 hypothetical prote
539 28 75.7 636 2 A84903 hypothetical prote
540 28 75.7 638 2 A33991 somatotropin recep

541	28	75.7	638	2	S12136	somatotropin recep	614	28	75.7	1529	2	T16779	hypothetical prote
542	28	75.7	638	2	B28176	somatotropin recep	615	28	75.7	1599	2	T15854	hypothetical prote
543	28	75.7	638	2	A33505	somatotropin recep	616	28	75.7	1733	2	S27939	tensin - chicken
544	28	75.7	643	2	S60710	alpha 1,2 mannosid	617	28	75.7	1744	2	A54970	tensin, cardiac mu
545	28	75.7	643	2	T23453	hypothetical prote	618	28	75.7	1792	2	A57075	tensin - chicken (
546	28	75.7	646	1	WZBEC8	68.6K capsid prote	619	28	75.7	1824	2	T07589	disease resistance
547	28	75.7	677	2	S33608	somatotropin-bindi	620	28	75.7	1976	2	I56555	sodium channel pro
548	28	75.7	699	2	C43674	US4 protein - huma	621	28	75.7	2133	2	T42763	coagulation factor
549	28	75.7	704	2	T24517	hypothetical prote	622	28	75.7	2403	2	A59386	sanko - human
550	28	75.7	705	2	T10723	NADPH-ferrihemopro	623	28	75.7	2411	2	A46299	tyrosine kinase su
551	28	75.7	712	2	T31523	hypothetical prote	624	28	75.7	2531	2	T31070	notch homolog - se
552	28	75.7	717	2	T27066	hypothetical prote	625	28	75.7	3010	1	GNWVCJ	genome polyprotein
553	28	75.7	738	2	S14270	alcohol dehydrogen	626	28	75.7	3157	2	B70969	probable PPE prote
554	28	75.7	770	2	S76883	nitric-oxide reduc	627	28	75.7	3624	2	AD0835	large repetitive p
555	28	75.7	780	2	T27941	hypothetical prote	628	28	75.7	3869	2	A48205	All-1 protein +GTE
556	28	75.7	786	2	E87565	conserved hypothet	629	28	75.7	4936	2	AH2515	hypothetical prote
557	28	75.7	791	2	A53691	diacylglycerol kin	630	28	75.7	4957	2	T03455	ALR protein - huma
558	28	75.7	793	2	S59067	penton long fiber	631	28	75.7	5262	2	T03454	ALR protein - huma
559	28	75.7	794	2	B87605	sensory box histid	632	27	73.0	57	2	S10782	salivary protein P
560	28	75.7	820	2	T00645	hypothetical prote	633	27	73.0	64	2	B41286	granulocyte-macrop
561	28	75.7	824	2	I50618	c-fps proto oncoge	634	27	73.0	68	2	E85918	hypothetical prote
562	28	75.7	824	2	T23923	hypothetical prote	635	27	73.0	79	1	PJHUSB	proline-rich pepti
563	28	75.7	834	2	JC8035	N-acetyltransferas	636	27	73.0	87	2	T00171	hypothetical prote
564	28	75.7	845	2	T00071	hypothetical prote	637	27	73.0	90	2	T18068	hypothetical prote
565	28	75.7	859	2	S51646	inscuteable protei	638	27	73.0	107	2	C72623	hypothetical prote
566	28	75.7	865	2	A47282	calcium-binding pr	639	27	73.0	109	2	S64309	hypothetical prote
567	28	75.7	866	1	JC4305	dynamain II - human	640	27	73.0	113	2	AE0176	conserved hypothet
568	28	75.7	868	2	A36878	dynamain 2 - rat	641	27	73.0	115	2	S75426	hypothetical prote
569	28	75.7	869	2	A25945	coagulation factor	642	27	73.0	116	2	A72685	hypothetical prote
570	28	75.7	870	2	B53165	dynamain II isoform	643	27	73.0	118	2	S28201	alpha-amylase inhi
571	28	75.7	870	2	A53165	dynamain II isoform	644	27	73.0	126	2	PC2273	cytochrome P450 pr
572	28	75.7	872	2	S33015	hypothetical prote	645	27	73.0	128	2	PC2260	cytochrome P450 pr
573	28	75.7	873	1	TVFVF	protein-tyrosine k	646	27	73.0	130	2	S70623	cytochrome P450 Cy
574	28	75.7	873	1	TVFVFS	protein-tyrosine k	647	27	73.0	130	2	S66474	cytochrome P450 (C
575	28	75.7	875	1	URXLA2	peptidylglycine mo	648	27	73.0	130	2	S26301	hypothetical prote
576	28	75.7	885	2	S67660	hypothetical prote	649	27	73.0	131	2	S66475	cytochrome P450 (C
577	28	75.7	906	2	T00039	hypothetical prote	650	27	73.0	132	2	T11239	hypothetical prote
578	28	75.7	922	2	T37256	metalloproteinase	651	27	73.0	134	2	JC5572	proline-rich prote
579	28	75.7	927	2	A48085	transcription fact	652	27	73.0	137	2	S00696	vitellogenin precu
580	28	75.7	935	2	S17855	peptidylglycine mo	653	27	73.0	138	2	D96715	protein FN2.10 [i
581	28	75.7	938	1	QQBE24	nuclear antigen EB	654	27	73.0	138	2	C84491	hypothetical prote
582	28	75.7	970	2	JQ0302	hypothetical 112K	655	27	73.0	141	2	S35611	cytochrome P450 4A
583	28	75.7	992	2	A31666	hypothetical prote	656	27	73.0	147	2	A36790	hypothetical prote
584	28	75.7	997	2	S44457	period clock prote	657	27	73.0	148	2	T02305	hypothetical prote
585	28	75.7	999	2	A34307	Ca2+-transporting	658	27	73.0	149	2	E95201	hypothetical prote
586	28	75.7	999	2	S72267	Ca2+-transporting	659	27	73.0	149	2	E98068	hypothetical prote
587	28	75.7	1013	2	T46422	hypothetical prote	660	27	73.0	152	2	JC4894	ubiquitin-protein
588	28	75.7	1029	2	H86179	hypothetical prote	661	27	73.0	159	2	T19624	hypothetical prote
589	28	75.7	1040	2	T29092	TSC-22 protein hom	662	27	73.0	162	2	AB0459	probable regulator
590	28	75.7	1050	2	JW0092	serine-threonine k	663	27	73.0	163	2	A71469	hypothetical prote
591	28	75.7	1053	2	S58883	calcium-channel ho	664	27	73.0	166	2	D87664	hypothetical prote
592	28	75.7	1087	2	C84263	transmembrane olig	665	27	73.0	169	2	T49618	2',3'-cyclic-nucle
593	28	75.7	1104	2	I38869	transcription fact	666	27	73.0	170	2	H75279	phenylacetic acid
594	28	75.7	1106	2	T44598	hypothetical prote	667	27	73.0	174	2	T45335	hypothetical prote
595	28	75.7	1106	2	T13938	gene shuttle craft	668	27	73.0	178	2	I50697	non-collagenous al
596	28	75.7	1110	2	T49091	gamma response I p	669	27	73.0	181	2	S37543	H+-transporting tw
597	28	75.7	1114	2	T30299	dynein heavy chain	670	27	73.0	181	2	T04643	hypothetical prote
598	28	75.7	1116	2	T38073	serine/threonine-p	671	27	73.0	181	2	T45581	hypothetical prote
599	28	75.7	1120	2	H88449	protein F54D8.1 [i	672	27	73.0	184	2	S48035	hypothetical prote
600	28	75.7	1166	2	T15628	hypothetical prote	673	27	73.0	187	2	H81347	hypothetical prote
601	28	75.7	1170	1	TSHUP1	thrombospondin 1 p	674	27	73.0	189	2	H81968	probable decarboxy
602	28	75.7	1206	2	S24407	formin isoform IV	675	27	73.0	190	2	B81023	3-octaprenyl-4-hyd
603	28	75.7	1240	2	T03097	CDO protein - huma	676	27	73.0	192	2	A55924	GTP-binding protei
604	28	75.7	1333	2	A37488	Ras guanine nucleo	677	27	73.0	195	2	T29395	hypothetical prote
605	28	75.7	1336	2	S25716	Ras guanine nucleo	678	27	73.0	197	2	F75436	conserved hypothet
606	28	75.7	1338	2	T30565	MAP kinase kinase	679	27	73.0	199	2	T16113	hypothetical prote
607	28	75.7	1409	2	T37188	presynaptic activi	680	27	73.0	203	2	G97380	phnH protein limpo
608	28	75.7	1464	2	T13716	bazooka gene prote	681	27	73.0	203	2	AF2598	conserved hypothet
609	28	75.7	1468	2	S11515	formin - mouse	682	27	73.0	204	2	G69761	phenylacrylic acid
610	28	75.7	1478	2	S20117	protein kinase BCK	683	27	73.0	205	2	S37804	hypothetical prote
611	28	75.7	1495	2	S60255	transcription co-r	684	27	73.0	216	2	S58652	hypothetical prote
612	28	75.7	1499	2	B59431	Rho GTPase activat	685	27	73.0	217	2	T24867	hypothetical prote
613	28	75.7	1503	2	T43166	alpha-2-macroglobu	686	27	73.0	218	2	E70944	probable lppi prot

687 27 73.0 221 2 D86600 SET domain protein
688 27 73.0 221 2 B81514 conserved hypothet
689 27 73.0 221 2 D72024 set domain protein
690 27 73.0 225 2 S55720 ribosomal protein
691 27 73.0 228 2 C70623 hypothetical prote
692 27 73.0 231 2 C70510 hypothetical prote
693 27 73.0 233 2 D86603 muramidase [import
694 27 73.0 233 2 A72022 muramidase (invasi
695 27 73.0 233 2 T28914 hypothetical prote
696 27 73.0 235 2 E84313 hypothetical prote
697 27 73.0 237 2 B70529 hypothetical prote
698 27 73.0 237 2 G72676 hypothetical prote
699 27 73.0 238 2 T11223 hypothetical prote
700 27 73.0 240 2 S53403 probable membrane
701 27 73.0 250 2 T36428 hypothetical prote
702 27 73.0 252 2 AD2558 hypothetical prote
703 27 73.0 252 2 T04739 hypothetical prote
704 27 73.0 254 2 AE2516 hypothetical prote
705 27 73.0 268 2 T15169 hypothetical prote
706 27 73.0 269 2 D88130 protein F10G7.6 [i
707 27 73.0 278 2 E82965 hypothetical prote
708 27 73.0 282 2 D95892 probable glycerol-
709 27 73.0 283 2 G69844 conserved hypothet
710 27 73.0 289 2 S74632 hypothetical prote
711 27 73.0 293 2 T45538 moaC protein [impo
712 27 73.0 294 1 S73832 MG237 homolog F10_
713 27 73.0 296 1 PDBYC REP2 protein - Yea
714 27 73.0 296 2 T47598 hypothetical prote
715 27 73.0 299 2 JQ1930 V protein - rinder
716 27 73.0 300 2 AH2955 hypothetical prote
717 27 73.0 300 2 E98327 hypothetical prote
718 27 73.0 301 2 E72289 oligopeptide ABC t
719 27 73.0 302 2 H96792 unknown protein F1
720 27 73.0 303 2 T46416 hypothetical prote
721 27 73.0 307 2 B87659 hydrolase, probabl
722 27 73.0 308 2 I46019 coatomer complex e
723 27 73.0 308 2 I48080 coatomer complex e
724 27 73.0 311 2 F86341 hypothetical prote
725 27 73.0 316 2 S16681 homeotic protein -
726 27 73.0 321 2 F84597 probable proline-r
727 27 73.0 324 2 A46631 lactose-binding le
728 27 73.0 325 2 D75555 conserved hypothet
729 27 73.0 329 2 B41344 lutropin-choriogon
730 27 73.0 329 2 D41344 lutropin-chorioi
731 27 73.0 331 2 B47236 zinc-finger protei
732 27 73.0 331 2 C41344 lutropin-choriogon
733 27 73.0 333 2 T23840 hypothetical prote
734 27 73.0 336 2 S29507 metavinculin - pig
735 27 73.0 342 2 B84220 hypothetical prote
736 27 73.0 344 2 A86294 hypothetical prote
737 27 73.0 349 2 D88431 protein M88.4 [imp
738 27 73.0 350 2 E71248 hypothetical prote
739 27 73.0 352 2 T06482 probable cell wall
740 27 73.0 352 2 C96643 hypothetical prote
741 27 73.0 356 2 I40212 cytochrome P450 BJ
742 27 73.0 362 2 C90674 hypothetical prote
743 27 73.0 362 2 C64758 yaha protein - Esc
744 27 73.0 363 2 B45900 complement C3d/Eps
745 27 73.0 365 2 F85524 hypothetical prote
746 27 73.0 366 2 C70907 hypothetical prote
747 27 73.0 367 2 H75440 alanine racemase -
748 27 73.0 368 2 T03694 cysteine proteinas
749 27 73.0 370 1 S55333 protein kinase pim
750 27 73.0 371 1 WMBEY0 dUTP diphosphatase
751 27 73.0 379 2 S31719 proline-rich prote
752 27 73.0 382 2 I39068 proline- arginine-
753 27 73.0 383 2 S76334 hypothetical prote
754 27 73.0 384 2 H70580 hypothetical prote
755 27 73.0 388 2 T27645 hypothetical prote
756 27 73.0 390 2 JE0212 hypothetical 44k p
757 27 73.0 390 2 JE0211 hypothetical 44k p
758 27 73.0 390 2 T49705 related to transcr
759 27 73.0 390 2 G01936 Abl binding protei

760 27 73.0 391 2 S54161 L-arginine-glycine
761 27 73.0 394 2 B75404 probable amidase -
762 27 73.0 395 2 G87481 N-acetylmuramoyl-L
763 27 73.0 396 1 S33310 endopeptidase I -
764 27 73.0 398 2 D83214 conserved hypothet
765 27 73.0 400 2 C70837 hypothetical prote
766 27 73.0 400 2 S06945 granulocyte-macrop
767 27 73.0 403 2 A53662 homeotic protein H
768 27 73.0 405 2 T10630 ethylene-regulated
769 27 73.0 409 2 S70704 carbon catabolite
770 27 73.0 410 2 B35670 protein-tyrosine k
771 27 73.0 412 2 T40899 probable proteinas
772 27 73.0 415 1 A34170 acrosin (EC 3.4.21
773 27 73.0 415 2 S12357 interleukin-5 rece
774 27 73.0 416 2 T28756 hypothetical prote
775 27 73.0 416 2 I38428 T-complex protein
776 27 73.0 416 2 A43561 homeotic protein m
777 27 73.0 417 2 T33376 hypothetical prote
778 27 73.0 419 2 T19871 hypothetical prote
779 27 73.0 420 2 H84182 hypothetical prote
780 27 73.0 423 2 S41734 glycine amidinotra
781 27 73.0 427 2 C83464 hypothetical prote
782 27 73.0 427 2 T23954 hypothetical prote
783 27 73.0 431 2 S47538 acrosin (EC 3.4.21
784 27 73.0 436 2 T22253 hypothetical prote
785 27 73.0 440 2 JC7807 Wiskott-Aldrich sy
786 27 73.0 448 2 T06076 proline-rich prote
787 27 73.0 451 2 B70805 hypothetical prote
788 27 73.0 459 2 G86264 F3F19 hypothetical
789 27 73.0 461 2 H84644 hypothetical prote
790 27 73.0 463 2 T01872 hypothetical prote
791 27 73.0 465 2 G02738 FREAC-4 - human
792 27 73.0 466 1 P2WL31 L2 protein - human
793 27 73.0 469 2 AI2833 UDP-N-acetylmuramo
794 27 73.0 469 2 E97611 hypothetical prote
795 27 73.0 470 2 S28303 hypothetical prote
796 27 73.0 472 1 P2WL58 L2 protein - human
797 27 73.0 473 1 P2WLHS minor capsid prote
798 27 73.0 476 2 AG1051 probable transport
799 27 73.0 478 2 T03548 mannitol 2-dehydro
800 27 73.0 479 2 T09365 cytochrome P450 ho
801 27 73.0 480 2 T48311 zinc finger protei
802 27 73.0 480 2 S52306 zinc finger protei
803 27 73.0 480 2 C86357 hypothetical prote
804 27 73.0 481 2 T16484 hypothetical prote
805 27 73.0 483 2 T06712 probable cytochrom
806 27 73.0 484 2 D65230 hypothetical 52.9
807 27 73.0 484 2 A98275 hypothetical prote
808 27 73.0 484 2 A86116 hypothetical prote
809 27 73.0 487 2 B70720 hypothetical prote
810 27 73.0 488 2 T42038 catalase (EC 1.11.
811 27 73.0 490 2 T00404 probable cytochrom
812 27 73.0 490 2 AE3433 mitochondrial proc
813 27 73.0 491 2 D95962 probable transcrip
814 27 73.0 492 2 T43064 CED-6 protein - Ca
815 27 73.0 496 1 S41192 cytochrome P450 4D
816 27 73.0 498 2 F84667 probable cytochrom
817 27 73.0 499 1 A40092 protein-tyrosine k
818 27 73.0 499 2 T49064 cytochrome P450-li
819 27 73.0 501 2 S76563 hypothetical prote
820 27 73.0 501 2 T28970 hypothetical prote
821 27 73.0 503 1 S66374 cytochrome P450 4M
822 27 73.0 503 2 T02357 cytochrome P450 ho
823 27 73.0 503 2 T48825 hypothetical prote
824 27 73.0 504 1 S23558 ubiquinol-cytochro
825 27 73.0 506 1 A40164 cytochrome P450 4B
826 27 73.0 507 1 A32966 cytochrome P450 4A
827 27 73.0 507 1 A43387 polymerase-associa
828 27 73.0 507 2 JC7883 cytochrome P450 4X
829 27 73.0 507 2 JQ1929 phosphoprotein - r
830 27 73.0 507 2 C71412 probable hydroxypr
831 27 73.0 509 1 O4RTLO laurate omega-hydr
832 27 73.0 509 2 S47553 cytochrome P450 Cy

833 27 73.0 510 1 A29368 prostaglandin omeg
834 27 73.0 510 1 A34160 laurate omega-hydr
835 27 73.0 510 2 T05939 cytochrome P450 mo
836 27 73.0 511 1 A34260 laurate omega-hydr
837 27 73.0 511 1 B34160 cytochrome P450 4A
838 27 73.0 511 1 B40164 cytochrome P450 4B
839 27 73.0 511 1 O4HUB1 cytochrome P450 4B
840 27 73.0 511 2 S66472 cytochrome P450 4B
841 27 73.0 511 2 T00864 cytochrome P450 ho
842 27 73.0 512 2 T00605 probable cytochrom
843 27 73.0 513 2 AI2555 hypothetical prote
844 27 73.0 514 2 T08014 cytochrome P450 Cy
845 27 73.0 514 2 D84584 probable protein p
846 27 73.0 515 2 I49154 calcitonin recepto
847 27 73.0 515 2 I60800 calcitonin recepto
848 27 73.0 516 2 T00514 cytochrome P450 ho
849 27 73.0 516 2 T09364 cytochrome P450 ho
850 27 73.0 519 2 I53015 fatty acid omega-h
851 27 73.0 519 2 JX0331 laurate omega-hydr
852 27 73.0 520 1 A46661 leukotriene B4 ome
853 27 73.0 520 2 S45702 leukotriene-B4 20-
854 27 73.0 523 2 C84753 hypothetical prote
855 27 73.0 526 2 T13687 hypothetical prote
856 27 73.0 527 2 A84645 probable cytochrom
857 27 73.0 531 1 SYNCYT tyrosine-trNA liga
858 27 73.0 533 1 S56652 calcium-dependent
859 27 73.0 536 2 S33569 protein-tyrosine k
860 27 73.0 538 2 T03262 cytochrome P450 -
861 27 73.0 547 2 G75364 DNA mismatch repai
862 27 73.0 552 2 A96756 hypothetical prote
863 27 73.0 570 2 T27407 hypothetical prote
864 27 73.0 576 2 T38293 hypothetical serin
865 27 73.0 577 2 H86414 F28N24.8 protein -
866 27 73.0 578 2 E85048 hypothetical prote
867 27 73.0 581 2 E91165 gamma-glutamyltran
868 27 73.0 581 2 E86011 gamma-glutamyltran
869 27 73.0 581 2 G86338 protein F2D10.4 li
870 27 73.0 582 2 A87316 hypothetical prote
871 27 73.0 583 2 A70729 hypothetical prote
872 27 73.0 584 2 H87304 hypothetical prote
873 27 73.0 589 2 F86549 CT365 hypothetical
874 27 73.0 589 2 D72074 conserved hypothet
875 27 73.0 589 2 T34596 probable membrane
876 27 73.0 591 2 I65981 fatty acid omega-h
877 27 73.0 598 2 T02795 probable membrane
878 27 73.0 599 2 JH0259 prostaglandin-endo
879 27 73.0 601 2 T49745 sensor histidine k
880 27 73.0 605 2 B87365 somatotropin recep
881 27 73.0 608 2 S32823 hypothetical prote
882 27 73.0 613 2 S27770 hypothetical prote
883 27 73.0 616 2 JQ1441 hypothetical 67K p
884 27 73.0 616 2 I38155 DNA-binding regula
885 27 73.0 632 2 E69407 NADH oxidase (noxB
886 27 73.0 632 2 G69306 NADH oxidase (noxB
887 27 73.0 634 2 F82623 potassium uptake p
888 27 73.0 637 2 A36427 lamin B receptor -
889 27 73.0 640 2 S67656 hypothetical prote
890 27 73.0 641 2 JC7331 gamma-glutamyltran
891 27 73.0 650 2 T04830 probable serine/th
892 27 73.0 671 2 F82960 hypothetical prote
893 27 73.0 674 2 S74506 ribonuclease E - S
894 27 73.0 676 2 A45900 complement C3b rec
895 27 73.0 678 2 B70913 probable penicilli
896 27 73.0 679 2 B75262 conserved hypothet
897 27 73.0 681 2 T17342 hypothetical prote
898 27 73.0 695 2 S51433 MDL1 protein - yea
899 27 73.0 696 2 A41344 lutropin-choriogon
900 27 73.0 699 1 QRHUUT lutropin-choriogon
901 27 73.0 699 2 D82798 phage-related term
902 27 73.0 702 2 JC7890 fructan beta-fruct
903 27 73.0 704 2 JC7932 phosphorylated car
904 27 73.0 708 2 F87245 penicillin-bonding
905 27 73.0 718 2 AI3420 penicillin-binding

906 27 73.0 730 2 S48813 hypothetical prote
907 27 73.0 742 2 T33514 hypothetical prote
908 27 73.0 743 2 B88551 protein T23G5.2 li
909 27 73.0 743 2 E71432 hypothetical prote
910 27 73.0 748 2 T49633 glucan 1,4-alpha-g
911 27 73.0 777 2 A35966 chemotaxis protein
912 27 73.0 794 2 C87437 penicillin-binding
913 27 73.0 834 2 S66258 glucosidase I - hu
914 27 73.0 838 2 I45557 eyeless, long form
915 27 73.0 859 2 T35785 probable beta-gluc
916 27 73.0 864 2 H85335 hypothetical prote
917 27 73.0 864 2 T04518 hypothetical prote
918 27 73.0 873 2 A47283 calphotin - fruit
919 27 73.0 896 2 T00389 hypothetical prote
920 27 73.0 903 2 T09143 alpha-glucosidase
921 27 73.0 913 2 S61580 probable membrane
922 27 73.0 929 2 JC6124 diacylglycerol kin
923 27 73.0 972 2 T50400 origin recognition
924 27 73.0 982 1 GNLJH2 pol polyprotein -
926 27 73.0 992 2 T46337 hypothetical prote
927 27 73.0 1004 2 A48821 hypothetical prote
928 27 73.0 1006 2 T00050 Wnt-5 protein - fr
929 27 73.0 1006 2 C86292 hypothetical prote
930 27 73.0 1014 2 S48235 CIN1 protein - yea
931 27 73.0 1016 2 D86308 translation initia
932 27 73.0 1025 1 A43526 complement C3d/Eps
933 27 73.0 1033 2 T37715 actin-interacting
934 27 73.0 1036 2 D84741 probable cellulose
935 27 73.0 1042 1 GBECE beta-galactosidase
936 27 73.0 1042 2 E85968 evolved beta-D-gal
937 27 73.0 1042 2 F91123 evolved beta-D-gal
938 27 73.0 1046 1 GNWEC genome polypotein
939 27 73.0 1048 2 A70592 hypothetical prote
940 27 73.0 1061 1 GNLJG4 HIV-1 retropepsin
941 27 73.0 1066 2 T10108 vinculin - mouse
942 27 73.0 1086 2 T33893 hypothetical prote
943 27 73.0 1089 2 C70522 probable mmpL8 pro
944 27 73.0 1102 2 T28666 protein kinase C-r
945 27 73.0 1107 1 JQ1658 genome polypotein
946 27 73.0 1134 1 A35955 meta-vinculin - hu
947 27 73.0 1135 1 A29997 meta-vinculin - ch
948 27 73.0 1151 2 G96805 hypothetical protei
949 27 73.0 1167 2 T34020 zinc finger protei
950 27 73.0 1186 2 T51793 hypothetical prote
951 27 73.0 1186 2 T33754 O/E-1-associated z
952 27 73.0 1199 2 A41939 G protein-coupled
953 27 73.0 1201 2 T29329 hypothetical prote
954 27 73.0 1202 2 PQ0440 polypotein - barl
955 27 73.0 1205 2 T27053 hypothetical prote
956 27 73.0 1259 2 A43425 Bravo/Nr-CAM cell
957 27 73.0 1262 2 T13353 protein stn-B - fr
958 27 73.0 1268 1 A39640 neural cell adhesi
959 27 73.0 1271 2 T24008 hypothetical prote
960 27 73.0 1280 2 T00365 hypothetical prote
961 27 73.0 1284 2 T40879 probable helicase
962 27 73.0 1299 2 AH2090 two-component hybr
963 27 73.0 1361 2 T30884 neural specific DN
964 27 73.0 1386 2 T49316 profilaggrin relat
965 27 73.0 1402 2 I46707 translation initia
966 27 73.0 1429 2 T13720 gene expanded prot
967 27 73.0 1443 2 S05979 steroid hormone re
968 27 73.0 1445 2 A59437 KIAA1204 protein l
969 27 73.0 1494 2 T14355 protein-tyrosine-p
970 27 73.0 1520 2 T00273 hypothetical prote
971 27 73.0 1522 2 T39371 transcription regu
972 27 73.0 1532 2 A26039 IGA-specific metal
973 27 73.0 1561 2 S61314 IGA-specific metal
974 27 73.0 1603 2 S17983 gene posterior sex
975 27 73.0 1693 2 AC3240 helicase, SNF2 fam
976 27 73.0 1759 2 T18868 myoblast city prot
977 27 73.0 1763 2 S16366 collagen alpha 2(I
978 27 73.0 1773 2 A81937 IGA-specific metal

979 27 73.0 1815 2 C81169 IgA-specific metal
980 27 73.0 1842 2 T43409 probable fatty-aci
981 27 73.0 1842 2 T38781 fatty acid synthas
982 27 73.0 1908 2 A86311 protein Fil3.14 [I
983 27 73.0 1941 2 T23979 hypothetical prote
984 27 73.0 1943 2 T23986 hypothetical prote
985 27 73.0 1952 2 T48814 hypothetical prote
986 27 73.0 2175 1 S03170 homeotic protein c
987 27 73.0 2301 2 T02323 nodulin-like prote
988 27 73.0 2424 2 I46480 calcium channel BI
989 27 73.0 2468 2 A83412 hypothetical prote
990 27 73.0 2523 2 F70846 probable PPE prote
991 27 73.0 2672 2 A48126 translation activa
992 27 73.0 2824 2 T22759 hypothetical prote
993 27 73.0 2897 2 B48666 cell proliferation
994 27 73.0 3122 2 T17202 DNA-directed DNA p
995 27 73.0 3256 2 A48666 cell proliferation
996 27 73.0 3386 1 GNWVDF genome polyprotein
997 27 73.0 3511 2 A59295 unconventional myo
998 27 73.0 3716 2 E70969 probable PPE prote
999 27 73.0 5175 2 T20992 hypothetical prote
1000 27 73.0 5198 2 T43290 hemicentin precurs

ALIGNMENTS

RESULT 1
B49502
protein-tyrosine-phosphatase (EC 3.1.3.48) receptor type 4E, splice form B precursor - f
C;Species: Drosophila melanogaster
C;Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: B49502
R;Oon, S.H.; Hong, A.; Yang, X.; Chia, W.
J. Biol. Chem. 268, 23964-23971, 1993
A;Title: Alternative splicing in a novel tyrosine phosphatase gene (DPTP4E) of Drosophil
A;Reference number: A49502; MUID:94043220; PMID:8226938
A;Accession: B49502
A;Molecule type: mRNA
A;Residues: 1-1615 <OON>
A;Cross-references: UNIPROT:Q9W4F5; UNIPARC:UPI0000177053; GB:L20894
C;Genetics:
A;Gene: FlyBase:Ptp4E
A;Cross-references: FlyBase:FBgn0004368
A;Introns: 1605/3
C;Superfamily: protein-tyrosine-phosphatase, receptor type 4E; fibronectin type III repe
C;Keywords: alternative splicing; phosphoprotein; phosphoric monoester hydrolase; recept
F;1254-1270/Domain: transmembrane #status predicted <TMN>
F;1271-1615/Domain: intracellular #status predicted <INT>
F;1353-1573/Domain: protein-tyrosine-phosphatase homology <PTP1>
F;1525/Active site: Cys (phosphocysteine intermediate) #status predicted
F;1531/Binding site: substrate phosphate (Arg) #status predicted

Query Match 100.0%; Score 37; DB 2; Length 1615;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
| | | | |
Db 1101 ILAPPVP 1107

RESULT 2
A49502
protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type 4E, splice form A precursor -
C;Species: Drosophila melanogaster
C;Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 31-Dec-2004
C;Accession: A49502
R;Oon, S.H.; Hong, A.; Yang, X.; Chia, W.
J. Biol. Chem. 268, 23964-23971, 1993
A;Title: Alternative splicing in a novel tyrosine phosphatase gene (DPTP4E) of Drosophil
A;Reference number: A49502; MUID:94043220; PMID:8226938
A;Accession: A49502

A;Molecule type: mRNA
A;Residues: 1-1767 <OON>
A;Cross-references: UNIPROT:Q9W4F5; UNIPROT:Q24495; UNIPARC:UPI0000177052; GB:L20894
A;Note: authors translated the codon ATA for residue 1715 as Leu
C;Genetics:
A;Gene: ptp4E
A;Cross-references: FlyBase:FBgn0004368
C;Superfamily: fibronectin type III repeat homology; protein-tyrosine-phosphatase homole
C;Keywords: alternative splicing; phosphoprotein; phosphoric monoester hydrolase; recept
F;1254-1270/Domain: transmembrane #status predicted <TMN>
F;1271-1767/Domain: intracellular #status predicted <INT>
F;1353-1573/Domain: protein-tyrosine-phosphatase homology <PTP1>
F;1525/Active site: Cys (phosphocysteine intermediate) #status predicted
F;1531/Binding site: substrate phosphate (Arg) #status predicted

Query Match 100.0%; Score 37; DB 2; Length 1767;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
| | | | |
Db 1101 ILAPPVP 1107

RESULT 3
F70595
hypothetical protein RV3213c - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: F70595
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: F70595
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-266 <COL>
A;Cross-references: UNIPROT:O05853; UNIPARC:UPI00000D4F2A; GB:Z95120; GB:AL123456; NID:g
A;Experimental source: strain H37RV
C;Genetics:
A;Gene: RV3213c
C;Superfamily: regulatory protein spo0J

Query Match 94.6%; Score 35; DB 2; Length 266;
Best Local Similarity 71.4%; Pred. No. 38;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
: | | | | :
Db 209 VLAPPIP 215

RESULT 4
T24591
hypothetical protein T06E4.7 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T24591
R;Lloyd, C.
submitted to the EMBL Data Library, April 1996
A;Reference number: Z19910
A;Accession: T24591
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-441 <WIL>
A;Cross-references: UNIPROT:Q22261; UNIPARC:UPI0000082B31; EMBL:Z70756; PIDN:CAA94793.1;
A;Experimental source: clone T06E4
C;Genetics:
A;Gene: CESP:T06E4.7

A;Map position: 5
A;Introns: 104/2; 317/2; 354/3; 404/3

Query Match 94.6%; Score 35; DB 2; Length 441;
Best Local Similarity 85.7%; Pred. No. 65;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
:|||||

Db 210 LLAPPVP 216

RESULT 5
D71658
3-octaprenyl-4-hydroxybenzoate carboxy-lyase (ubiX) RP541 - Rickettsia prowazekii
C;Species: Rickettsia prowazekii
C;Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 09-Jul-2004
C;Accession: D71658
R;Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmark, U.
Nature 396, 133-140, 1998
A;Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A;Reference number: A71630; MUID:99039499; PMID:9823893
A;Accession: D71658
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-189 <AND>
A;Cross-references: UNIPROT:Q9ZD09; UNIPARC:UPI0000131197; GB:AJ235272; GB:AJ235269; NID
A;Experimental source: strain Madrid E
C;Genetics:
A;Gene: ubiX; RP541
C;Superfamily: dedF protein

Query Match 91.9%; Score 34; DB 2; Length 189;
Best Local Similarity 71.4%; Pred. No. 40;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
|:|||||

Db 146 ILAPPVP 152

RESULT 6
S75284
chemotaxis protein cheA homolog - Synechocystis sp. (strain PCC 6803)
N;Alternate names: protein sll1296
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C;Accession: S75284
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S75284
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-924 <KAN>
A;Cross-references: UNIPROT:P73172; UNIPARC:UPI00000D34AC; EMBL:D90904; GB:AB001339; NID
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Genetics:
A;Gene: cheA
C;Keywords: chemotaxis; phosphoprotein
F;802-914/Domain: response regulator homology <RRH>
F;851/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match 91.9%; Score 34; DB 2; Length 924;
Best Local Similarity 71.4%; Pred. No. 2.2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
::|||||

Db 279 VIAPPVP 285

RESULT 7
AE1448
hypothetical protein lin0124 [imported] - Listeria innocua (strain Clip11262)
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: AE1448
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AE1448
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-113 <GLA>
A;Cross-references: UNIPROT:Q925X3; UNIPARC:UPI00001392F0; GB:AL592022; PIDN:CAC95357.1;
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: lin0124

Query Match 89.2%; Score 33; DB 2; Length 113;
Best Local Similarity 71.4%; Pred. No. 35;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
::|||||

Db 36 LIAPPVP 42

RESULT 8
AE1729
hypothetical protein lin2378 [imported] - Listeria innocua (strain Clip11262)
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: AE1729
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AE1729
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-113 <GLA>
A;Cross-references: UNIPROT:Q925X3; UNIPARC:UPI00001392F0; GB:AL592022; PIDN:CAC97605.1;
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: lin2378

Query Match 89.2%; Score 33; DB 2; Length 113;
Best Local Similarity 71.4%; Pred. No. 35;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
::|||||

Db 36 LIAPPVP 42

RESULT 9
AG1752
hypothetical protein lin2564 [imported] - Listeria innocua (strain Clip11262)
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: AG1752

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeckel
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of *Listeria* species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AG1752
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-113 <GLA>
A;Cross-references: UNIPROT:Q925X3; UNIPARC:UPI00001392F0; GB:AL592022; PIDN:CAC97791.1;
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: lin2564

Query Match 89.2%; Score 33; DB 2; Length 113;
Best Local Similarity 71.4%; Pred. No. 35;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
:|||||
Db 36 LIAPPVP 42

RESULT 10
S41002
hypothetical protein T05G5.2 - *Caenorhabditis elegans*
C;Species: *Caenorhabditis elegans*
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C;Accession: S41002
R;Thomas, K.
submitted to the EMBL Data Library, October 1993
A;Reference number: S41001
A;Accession: S41002
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-205 <THO>
A;Cross-references: UNIPROT:P34555; UNIPARC:UPI0000061154; EMBL:Z27079; NID:G414641; PID
C;Genetics:
A;Introns: 21/3

Query Match 89.2%; Score 33; DB 2; Length 205;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAPPVP 7
|||||
Db 139 LAPPVP 144

RESULT 11
T35712
integral membrane protein - *Streptomyces coelicolor*
C;Species: *Streptomyces coelicolor*
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C;Accession: T35712
R;Murphy, L.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, January 1998
A;Reference number: Z21548
A;Accession: T35712
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-465 <MUR>
A;Cross-references: UNIPROT:O54176; UNIPARC:UPI00000DABB3; EMBL:AL021411; PIDN:CAA16204.
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SC0EDB:SC7H1.17
C;Superfamily: multidrug-efflux transporter

Query Match 89.2%; Score 33; DB 2; Length 465;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 LAPPVP 7
|||||
Db 5 LAPPVP 10

RESULT 12
S76481
hypothetical protein - *Synechocystis* sp. (strain PCC 6803)
C;Species: *Synechocystis* sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C;Accession: S76481
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S76481
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-538 <KAN>
A;Cross-references: UNIPROT:P74506; UNIPARC:UPI00000C104B; EMBL:D90915; GB:AB001339; NID
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 89.2%; Score 33; DB 2; Length 538;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
|||||
Db 426 IFAPPVP 432

RESULT 13
AF1927
hypothetical protein all0969 [imported] - *Nostoc* sp. (strain PCC 7120)
C;Species: *Nostoc* sp. PCC 7120
A;Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AF1927
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Ana*
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AF1927
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-543 <KUR>
A;Cross-references: UNIPROT:Q8YV82; UNIPARC:UPI00000CDF2B; GB:BA000019; PIDN:BA072926.1;
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all0969

Query Match 89.2%; Score 33; DB 2; Length 543;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
|||||
Db 428 IFAPPVP 434

RESULT 14
T13994
envelope protein - fruit fly (*Drosophila melanogaster*) ZAM retrovirus-like element
C;Species: *Drosophila melanogaster*
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T13994
R;Leblanc, P.; Desset, S.; Dastugue, B.; Vaury, C.

EMBO J. 16, 7521-7531, 1997
A;Title: Invertebrate retroviruses: ZAM a new candidate in Drosophila melanogaster.
A;Reference number: Z17849; MUID:98070328; PMID:9405380
A;Accession: T13994
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-551 <LEB>
A;Cross-references: UNIPROT:O46113; UNIPARC:UPI000007A36F; EMBL:AJ000387; NID:g2791286;
C;Genetics:
A;Gene: env
A;Cross-references: FlyBase:FBgn0023131
A;Mobile element: ZAM retrovirus-like element
A;Introns: 13/2

Query Match 89.2%; Score 33; DB 2; Length 551;
Best Local Similarity 71.4%; Pred. No. 1.9e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
|:||||:
Db 520 IVAPPIP 526

RESULT 15
T08877
Modin - Podospora anserina
C;Species: Podospora anserina
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C;Accession: T08877
R;Barreau, C.; Iskandar, M.; Loubradou, G.; Levallois, V.; Begueret, J.
Genetics 149, 915-926, 1998
A;Title: The mod-A suppressor of nonallelic heterokaryon incompatibility in Podospora an
A;Reference number: Z16508; MUID:98278809; PMID:9611202
A;Accession: T08877
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-687 <BAR>
A;Cross-references: UNIPROT:O59895; UNIPARC:UPI000006C81B; EMBL:AF025289; NID:g3115380;
A;Experimental source: strain het-c1 het-e1 het-d1 mod-B1
C;Genetics:
A;Gene: mod-A
A;Introns: 358/1
C;Superfamily: Podospora anserina Modin

Query Match 89.2%; Score 33; DB 2; Length 687;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAPPVP 7
|||||
Db 599 LAPPVP 604

RESULT 16
F87708
cell division protein FtsK, probable [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: F87708
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: F87708
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-819 <STO>
A;Cross-references: UNIPROT:Q9A262; UNIPARC:UPI00000C7B8B; GB:AE005673; NID:g13425470; E
C;Genetics:
A;Gene: CC3704

Query Match 89.2%; Score 33; DB 2; Length 819;
Best Local Similarity 85.7%; Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
|||||
Db 811 ILAPPTP 817

RESULT 17
A41401
mineralocorticoid receptor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 05-Oct-2004
C;Accession: A41401
R;Patel, P.D.; Sherman, T.G.; Goldman, D.J.; Watson, S.J.
Mol. Endocrinol. 3, 1877-1885, 1989
A;Title: Molecular cloning of a mineralocorticoid (type I) receptor complementary DNA fr
A;Reference number: A41401; MUID:90114194; PMID:2558305
A;Accession: A41401
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-981 <PAT>
A;Cross-references: UNIPROT:P22199; UNIPARC:UPI000012EDE6; GB:M36074; NID:g205340; PIDN.
C;Superfamily: Mineralocorticoid receptor (MR); erba transforming protein homology
C;Keywords: DNA binding; transcription regulation; zinc finger
F;602-877/Domain: erba transforming protein homology <ERBA>
F;604-624/Region: zinc finger
F;640-664/Region: zinc finger

Query Match 89.2%; Score 33; DB 2; Length 981;
Best Local Similarity 85.7%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
|||||
Db 473 ILGPPVP 479

RESULT 18
A29513
mineralocorticoid receptor - human
N;Alternate names: aldosterone receptor
C;Species: Homo sapiens (man)
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 05-Oct-2004
C;Accession: A29513
R;Arriza, J.L.; Weinberger, C.; Cerelli, G.; Glaser, T.M.; Handelin, B.L.; Housman, D.E.
Science 237, 268-275, 1987
A;Title: Cloning of human mineralocorticoid receptor complementary DNA: structural and f
A;Reference number: A29513; MUID:87263386; PMID:3037703
A;Accession: A29513
A;Molecule type: mRNA
A;Residues: 1-984 <ARR>
A;Cross-references: UNIPROT:P08235; UNIPARC:UPI000012EDE5; GB:M16801; NID:g187460; PIDN:
C;Genetics:
A;Gene: GDB:MLR
A;Cross-references: GDB:120188; OMIM:264350
A;Map position: 4q31-4q31
C;Superfamily: Mineralocorticoid receptor (MR); erba transforming protein homology
C;Keywords: DNA binding; transcription regulation; zinc finger
F;601-880/Domain: erba transforming protein homology <ERBA>
F;603-623/Region: zinc finger
F;639-663/Region: zinc finger

Query Match 89.2%; Score 33; DB 2; Length 984;
Best Local Similarity 85.7%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
|||||
Db 472 ILGPPVP 478

RESULT 19
T21913
hypothetical protein F44G4.8 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T21913; T22211
R/Thomas, K.
submitted to the EMBL Data Library, September 1995
A/Reference number: Z19486
A/Accession: T21913
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1367 <WIL>
A/Cross-references: UNIPROT:Q20120; UNIPARC:UPI0000080BD3; EMBL:Z54218; PIDN:CAA90958.1;
A/Experimental source: clone F37B12
R/Sims, M.
submitted to the EMBL Data Library, June 1995
A/Reference number: Z19530
A/Accession: T22211
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1367 <W12>
A/Cross-references: UNIPARC:UPI0000080BD3; EMBL:Z49910; PIDN:CAA90125.1; GSPDB:GN000020;
A/Experimental source: clone F44G4
C/Genetics:
A/Gene: CESP:F44G4.8
A/Map position: 2
A/Introns: 76/2; 109/2; 163/3; 229/1; 276/3; 329/3; 381/2; 430/3; 648/1; 780/2; 805/3; 805/3;
Query Match 89.2%; Score 33; DB 2; Length 1367;
Best Local Similarity 71.4%; Pred. No. 5.1e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ILAPPVP 7
Db 740 LMAPPVP 746
RESULT 20
I38346
elastic titin - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 09-Jul-2004
C/Accession: I38346
R/Labeit, S.; Kolmerer, B.
Science 270, 293-296, 1995
A/Title: Titins: giant proteins in charge of muscle ultrastructure and elasticity.
A/Reference number: A57430; MUID:96026330; PMID:7569978
A/Accession: I38346
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-7962 <RES>
A/Cross-references: UNIPROT:Q10465; UNIPARC:UPI000011010C; EMBL:X90569; NID:g1017426; P
C/Genetics:
A/Gene: GDB:TTN
A/Cross-references: GDB:127867; OMIM:188840
A/Map position: 2q31-2q31
Query Match 89.2%; Score 33; DB 2; Length 7962;
Best Local Similarity 71.4%; Pred. No. 3.3e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ILAPPVP 7
Db 5625 VVAPPVP 5631
RESULT 21
A97800
hypothetical protein ubiX [imported] - Rickettsia conorii (strain Malish 7)
C/Species: Rickettsia conorii
C/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C/Accession: A97800

R/Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Ro
Science 293, 2093-2098, 2001
A/Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A/Reference number: A97700; MUID:21442074; PMID:11557893
A/Accession: A97800
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-191 <KUR>
A/Cross-references: UNIPROT:Q92HH0; UNIPARC:UPI000000CBEE1; GB:AE006914; PIDN:AAL03339.1;
C/Genetics:
A/Gene: ubiX
C/Superfamily: dedF protein
Query Match 86.5%; Score 32; DB 2; Length 191;
Best Local Similarity 71.4%; Pred. No. 93;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ILAPPVP 7
Db 146 IISPPVP 152
RESULT 22
H72585
hypothetical protein APE1157 - Aeropyrum pernix (strain K1)
C/Species: Aeropyrum pernix
C/Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C/Accession: H72585
R/Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A/Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A/Reference number: A72450; MUID:99310339; PMID:10382966
A/Accession: H72585
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-210 <KAW>
A/Cross-references: UNIPROT:Q9YCV5; UNIPARC:UPI000005DE2B; DDBJ:AP000061; NID:GS104821;
A/Experimental source: strain K1
C/Genetics:
A/Gene: APE1157
Query Match 86.5%; Score 32; DB 2; Length 210;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 2 LAPPVP 7
Db 95 LAPPVP 100
RESULT 23
T10399
hypothetical protein 130 - Orgyia pseudotsugata nuclear polyhedrosis virus
C/Species: Orgyia pseudotsugata nuclear polyhedrosis virus, OpMNPV
C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C/Accession: T10399
R/Ahrens, C.A.; Russell, R.R.; Funk, C.J.; Evans, J.; Harwood, S.; Rohrmann, G.F.
Virology 229, 381-399, 1997
A/Title: The sequence of the Orgyia pseudotsugata multinucleocapsid nuclear polyhedrosis
A/Reference number: Z17011; MUID:97271300; PMID:9126251
A/Accession: T10399
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-228 <AHR>
A/Cross-references: UNIPARC:UPI000006186E; EMBL:U75930; NID:g2934903; PIDN:AAC59129.1; P
C/Superfamily: OpMNPV hypothetical protein 130
Query Match 86.5%; Score 32; DB 2; Length 228;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 2 LAPPVP 7

Db	: 218 LAPPVP 223	
RESULT 24		
D30857	hypothetical protein 3 - Orgyia pseudotsugata multicapsid nuclear polyhedrosis virus	
C:Species:	Orgyia pseudotsugata multicapsid nuclear polyhedrosis virus, OpMNPV	
C:Date:	31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004	
C:Accession:	D30857; JT0431	
R:Gombart, A.F.; Blissard, G.W.; Rohrmann, G.F.	submitted to GenBank, February 1989	
A:Reference number:	A30857	
A:Accession:	D30857	
A:Molecule type:	DNA	
A:Residues:	1-228 <GOM>	
A:Cross-references:	UNIPROT:P24080; UNIPARC:UPI000006186E; GB:D13929; GB:D00508; NID:g22	
C:Genetics:		
A:Map position:	84.45-87.70	
C:Superfamily:	OpMNPV hypothetical protein 130	
Query Match	86.5%;	Score 32; DB 2; Length 228;
Best Local Similarity	83.3%;	Pred. No. 1.1e+02;
Matches	5; Conservative	1; Mismatches 0; Indels 0; Gaps 0;
QY	2 LAPPVP 7	
Db	: 218 LAPPVP 223	
RESULT 25		
E86433	protein T17H7.3 [imported] - Arabidopsis thaliana	
C:Species:	Arabidopsis thaliana (mouse-ear cress)	
C:Date:	02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004	
C:Accession:	E86433	
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Huizar, L. ansen, N.F.; Hughes, B.; Huizar, L. Nature 408, 816-820, 2000		
A:Authors:	Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C. C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marzialli, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.	
A:Authors:	Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.	
A:Title:	Sequence and analysis of chromosome 1 of the plant Arabidopsis.	
A:Reference number:	A86141; MUID:21016719; PMID:11130712	
A:Accession:	E86433	
A>Status:	preliminary	
A:Molecule type:	DNA	
A:Residues:	1-481 <STO>	
A:Cross-references:	UNIPROT:Q9SY17; UNIPARC:UPI00000A0CD1; GB:AE005172; NID:g4926818; PI	
C:Genetics:		
A:Gene:	T17H7.3	
A:Map position:	1	
Query Match	86.5%;	Score 32; DB 2; Length 481;
Best Local Similarity	71.4%;	Pred. No. 2.5e+02;
Matches	5; Conservative	2; Mismatches 0; Indels 0; Gaps 0;
QY	1 ILAPPVP 7	
Db	: 173 LLAPLP 179	
RESULT 26		
T24393	hypothetical protein T03E6.8 - Caenorhabditis elegans	
C:Species:	Caenorhabditis elegans	
C:Date:	15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004	
C:Accession:	T24393	
R:Lloyd, C.	submitted to the EMBL Data Library, March 1997	

A:Reference number:	219885	
A:Accession:	T24393	
A>Status:	preliminary; translated from GB/EMBL/DBJ	
A:Molecule type:	DNA	
A:Residues:	1-581 <WIL>	
A:Cross-references:	UNIPROT:O45740; UNIPARC:UPI000016427C; EMBL:Z92812; PIDN:CAB07281.1;	
A:Experimental source:	clone T03E6	
C:Genetics:		
A:Gene:	CESP:T03E6.8	
A:Map position:	5	
A:Introns:	60/2; 113/1; 182/3; 222/3; 249/1; 305/2; 354/1; 540/1; 566/2	
Query Match	86.5%;	Score 32; DB 2; Length 581;
Best Local Similarity	85.7%;	Pred. No. 3.1e+02;
Matches	6; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
QY	1 ILAPPVP 7	
Db	 320 ILEPPVP 326	
RESULT 27		
C86152	T7I23.2 protein - Arabidopsis thaliana	
C:Species:	Arabidopsis thaliana (mouse-ear cress)	
C:Date:	02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004	
C:Accession:	C86152	
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K. ansen, N.F.; Hughes, B.; Huizar, L. Nature 408, 816-820, 2000		
A:Authors:	Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C. C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marzialli, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.	
A:Authors:	Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.	
A:Title:	Sequence and analysis of chromosome 1 of the plant Arabidopsis.	
A:Reference number:	A86141; MUID:21016719; PMID:11130712	
A:Accession:	C86152	
A>Status:	preliminary	
A:Molecule type:	DNA	
A:Residues:	1-642 <STO>	
A:Cross-references:	UNIPROT:O23673; UNIPARC:UPI000009D3C0; GB:AE005172; NID:g2317902; PI	
C:Genetics:		
A:Map position:	1	
Query Match	86.5%;	Score 32; DB 2; Length 642;
Best Local Similarity	57.1%;	Pred. No. 3.4e+02;
Matches	4; Conservative	3; Mismatches 0; Indels 0; Gaps 0;
QY	1 ILAPPVP 7	
Db	: 186 LIAPPVP 192	
RESULT 28		
A96807	hypothetical protein T32E8.5 [imported] - Arabidopsis thaliana	
C:Species:	Arabidopsis thaliana (mouse-ear cress)	
C:Date:	02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004	
C:Accession:	A96807	
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K. ansen, N.F.; Hughes, B.; Huizar, L. Nature 408, 816-820, 2000		
A:Authors:	Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C. C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marzialli, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.	
A:Authors:	Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.	
A:Title:	Sequence and analysis of chromosome 1 of the plant Arabidopsis.	
A:Reference number:	A86141; MUID:21016719; PMID:11130712	
A:Accession:	A96807	

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-720 <STO>
A;Cross-references: UNIPROT:Q9CA22; UNIPARC:UPI000009F53F; GB:AE005173; NID:g6437532; PID:G3327097;
C;Genetics:
A;Gene: T32E8.5
A;Map position: 1

Query Match 86.5%; Score 32; DB 2; Length 720;
Best Local Similarity 83.3%; Pred. No. 3.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAPPVP 7
|||:|
Db 636 LAPPIP 641

RESULT 29
T01619
hypothetical protein At2g18910 [imported] - Arabidopsis thaliana
N;Alternate names: hypothetical protein F19F24.11
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 09-Jul-2004
C;Accession: T01619; B84570
R;Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul, M.; Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: B84570
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-753 <ROU>
A;Cross-references: UNIPROT:O64621; UNIPARC:UPI000017AF77; EMBL:AC003673; NID:g3004543;
A;Experimental source: cultivar Columbia
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: B84570
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-753 <STO>
A;Cross-references: UNIPARC:UPI000017AF77; GB:AE002093; NID:g3004565; PIDN:AAC09038.1; G3327097;
C;Genetics:
A;Gene: At2g18910; F19F24.11
A;Map position: 2
A;Introns: 78/2; 606/1; 705/1; 737/3

Query Match 86.5%; Score 32; DB 2; Length 753;
Best Local Similarity 83.3%; Pred. No. 4.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAPPVP 7
|||:|
Db 388 LAPPIP 393

RESULT 30
T00377
KIAA0642 protein - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
C;Accession: T00377
R;Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N. DNA Res. 5, 169-176, 1998
A;Title: Prediction of the coding sequences of unidentified human genes. X. The complete
A;Reference number: Z14142; MUID:98403880; PMID:9734811
A;Accession: T00377
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: mRNA
A;Residues: 1-1069 <ISH>

A;Cross-references: UNIPROT:O75137; UNIPARC:UPI000017C289; EMBL:AB014542; NID:g3327097;
C;Genetics:
A;Gene: KIAA0542

Query Match 86.5%; Score 32; DB 2; Length 1069;
Best Local Similarity 85.7%; Pred. No. 5.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
|||:|
Db 220 ILPPVP 226

RESULT 31
D89451
protein T04G9.1 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: D89451
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/projects/C_ele
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: D89451
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1610 <STO>
A;Cross-references: UNIPROT:Q22173; UNIPARC:UPI000007A764; GB:chr_X; PIDN:AAA82463.1; PI
C;Genetics:
A;Gene: T04G9.1
A;Map position: X

Query Match 86.5%; Score 32; DB 2; Length 1610;
Best Local Similarity 85.7%; Pred. No. 9.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
|||:|
Db 43 ILEPPVP 49

RESULT 32
A75613
hypothetical protein DRA0166 - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: A75613
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: A75613
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1626 <WHI>
A;Cross-references: UNIPROT:Q9RYV7; UNIPARC:UPI00000C164F; GB:AE001862; NID:G3327097;
A;Experimental source: strain R1
C;Genetics:
A;Gene: DRA0166
A;Map position: 2
C;Superfamily: Deinococcus radiodurans hypothetical protein DRA0166

Query Match 86.5%; Score 32; DB 2; Length 1626;
Best Local Similarity 71.4%; Pred. No. 9.3e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
:| |||
Db 752 VLGPVP 758

RESULT 33
T00390
KIAA0614 protein - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
C;Accession: T00390; T17254
R;Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.
DNA Res. 5, 169-176, 1998
A;Title: Prediction of the coding sequences of unidentified human genes. X. The complete
A;Reference number: Z14142; MUID:98403880; PMID:9734811
A;Accession: T00390
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1630 <ISH>
A;Cross-references: UNIPROT:Q9UFT6; UNIPARC:UPI000003B446; EMBL:AB014514; NID:g3327041;
R;Koehrer, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, September 1999
A;Reference number: Z18722
A;Accession: T17254
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1515-1630 <KOE>
A;Cross-references: UNIPARC:UPI0000073F07; EMBL:AL117469
A;Experimental source: adult uterus; clone DKFzp586O1022
C;Genetics:
A;Note: KIAA0614; DKFzp586O1022.1

Query Match 86.5%; Score 32; DB 2; Length 1630;
Best Local Similarity 71.4%; Pred. No. 9.3e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
:|||||
Db 122 LLAPPLP 128

RESULT 34
T24089
hypothetical protein R09E10.7 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T24089
R;Matthews, L.
submitted to the EMBL Data Library, March 1996
A;Reference number: Z19839
A;Accession: T24089
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1791 <WIL>
A;Cross-references: UNIPROT:Q21875; UNIPARC:UPI000007D7E4; EMBL:Z70287; PIDN:CAA94301.1;
A;Experimental source: clone R09E10
C;Genetics:
A;Gene: CESP:R09E10.7
A;Map position: 4
A;Introns: 78/2; 94/3; 250/2; 350/2; 674/3; 1079/2; 1155/2; 1450/3; 1593/3; 1642/2; 1690

Query Match 86.5%; Score 32; DB 2; Length 1791;
Best Local Similarity 83.3%; Pred. No. 1e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAPPVP 7
|||||
Db 592 LAPPIP 597

RESULT 35
A59266
unconventional myosin-15 - human
C;Species: Homo sapiens (man)
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C;Accession: A59266

R;Liang, Y.; Wang, A.; Belyantseva, I.A.; Anderson, D.W.; Probst, F.J.; Barber, T.D.; Mi
an, T.B.; Fridell, R.A.
Genomics 61, 243-258, 1999
A;Title: Characterization of the human and mouse unconventional myosin XV genes responsi
A;Reference number: A59266; MUID:20021762; PMID:10552926
A;Accession: A59266
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-3530 <LIA>
A;Cross-references: UNIPROT:Q9UKN7; UNIPARC:UPI000012FABF; GB:AF144094; NID:g6224682; PI
F;1225-1887/Domain: myosin motor domain homology <MMO>

Query Match 86.5%; Score 32; DB 2; Length 3530;
Best Local Similarity 71.4%; Pred. No. 2.1e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
:|||||
Db 660 LLSPPVP 666

RESULT 36
H87656
3-octaprenyl-4-hydroxybenzoate carboxy-lyase [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: H87656
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: H87656
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-162 <STO>
A;Cross-references: UNIPROT:Q9A3B4; UNIPARC:UPI00000C7A0D; GB:AE005673; NID:g13424982; P
C;Genetics:
A;Gene: CC3290

Query Match 83.8%; Score 31; DB 2; Length 162;
Best Local Similarity 57.1%; Pred. No. 1.2e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
:|||||
Db 109 VIAPPLP 115

RESULT 37
H89164
protein T06E4.8 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: H89164
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: H89164
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-198 <STO>
A;Cross-references: UNIPROT:Q22262; UNIPARC:UPI000017A685; GB:chr_V; PIDN:CAA94794.1; PI
C;Genetics:
A;Gene: T06E4.8
A;Map position: 5

Query Match 83.8%; Score 31; DB 2; Length 198;
Best Local Similarity 71.4%; Pred. No. 1.5e+02;

Matches	5;	Conservative	1;	Mismatches	1;	Indels	0;	Gaps	0;
Qy	1	ILAPPVP 7	:						
Db	103	LLAPPAP 109							
RESULT 38									
T24595									
hypothetical protein T06E4.8 - Caenorhabditis elegans									
C;Species: Caenorhabditis elegans									
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004									
C;Accession: T24595									
R;Lloyd, C.									
submitted to the EMBL Data Library, April 1996									
A;Reference number: Z19910									
A;Accession: T24595									
A;Status: preliminary; translated from GB/EMBL/DBDJ									
A;Molecule type: DNA									
A;Residues: 1-199 <WIL>									
A;Cross-references: UNIPROT:Q22262; UNIPARC:UPI0000081907; EMBL:Z70756; PIDN:CAA94794.2;									
A;Experimental source: clone T06E4									
C;Genetics:									
A;Gene: CESP:T06E4.8									
A;Map position: 5									
A;Introns: 21/2, 198/2									
Query Match 83.8%; Score 31; DB 2; Length 199;									
Best Local Similarity 71.4%; Pred. No. 1.5e+02;									
Matches	5;	Conservative	1;	Mismatches	1;	Indels	0;	Gaps	0;
Qy	1	ILAPPVP 7	:						
Db	103	LLAPPAP 109							
RESULT 39									
T51884									
hypothetical protein DKFZp547K054.1 - human (fragment)									
C;Species: Homo sapiens (man)									
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 09-Jul-2004									
C;Accession: T51884									
R;Bloeker, H.; Boecher, M.; Brandt, P.; Mewes, H.W.; Weil, B.; Wiemann, S.									
submitted to the Protein Sequence Database, July 2000									
A;Reference number: Z25853									
A;Accession: T51884									
A;Status: preliminary									
A;Molecule type: mRNA									
A;Residues: 1-205 <AAA>									
A;Cross-references: UNIPROT:Q9NPT3; UNIPARC:UPI000006E1B0; EMBL:AL390175									
A;Experimental source: fetal brain; clone DKFZp547K054									
C;Genetics:									
A;Note: DKFZp547K054.1									
Query Match 83.8%; Score 31; DB 2; Length 205;									
Best Local Similarity 71.4%; Pred. No. 1.5e+02;									
Matches	5;	Conservative	1;	Mismatches	1;	Indels	0;	Gaps	0;
Qy	1	ILAPPVP 7	:						
Db	55	LLCPPVP 61							
RESULT 40									
G95858									
probable 3-octaprenyl-4-hydroxybenzoate carboxy-lyase protein (EC 4.1.1.-) [imported] -									
C;Species: Sinorhizobium meliloti									
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004									
C;Accession: G95858									
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan									
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001									
A;Title: The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing endo									
A;Reference number: A95842; MUID:21396508; PMID:11481431									

A;Accession: G95858
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-210 <KUR>
A;Cross-references: UNIPROT:Q92X27; UNIPARC:UPI00000CB414; GB:AL591985; PIDN:CAC48535.1;
A;Experimental source: strain 1021, megaplasmid pSymb
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, P.;
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.P.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMB20135
A;Genome: plasmid
C;Superfamily: dedf protein
C;Keywords: carbon-carbon lyase; carboxy-lyase

Query Match	83.8%;	Score 31;	DB 2;	Length 210;
Best Local Similarity	71.4%;	Pred. No. 1.6e+02;		
Matches	5;	Conservative 1;	Mismatches 1;	Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
|: ||||
Db 150 IICPPVP 156

RESULT 41
T24594
hypothetical protein T06E4.9 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T24594
R;Lloyd, C.
submitted to the EMBL Data Library, April 1996
A;Reference number: Z19910
A;Accession: T24594
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-212 <WIL>
A;Cross-references: UNIPROT:Q22263; UNIPARC:UPI00000751FF; EMBL:Z70756; PIDN:CAA94795.2;
A;Experimental source: clone T06E4
C;Genetics:
A;Gene: CESP:T06E4.9
A;Map position: 5
A;Introns: 34/2; 211/2

Query Match	83.8%;	Score 31;	DB 2;	Length 212;
Best Local Similarity	71.4%;	Pred. No. 1.6e+02;		
Matches	5;	Conservative 1;	Mismatches 1;	Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
|: ||||
Db 116 LLAPPAP 122

RESULT 42
G89164
protein T06E4.9 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: G89164
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/projects/C_ele
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: G89164
A;Status: preliminary

A;Molecule type: DNA
A;Residues: 1-213 <STO>
A;Cross-references: UNIPROT:Q22263; UNIPARC:UPI000017A684; GB:chr_V; PIDN:CAA94795.1; PIDN:10-632-388-296
C;Genetics:
A;Gene: T06E4.9
A;Map position: 5

Query Match 83.8%; Score 31; DB 2; Length 213;
Best Local Similarity 71.4%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
:|||||
Db 116 LLAPPAP 122

RESULT 43
F95271
hypothetical protein SMA0148 [imported] - Sinorhizobium meliloti (strain 1021) magaplasma
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: F95271
R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bower, J.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C. Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti
A;Reference number: A95262; MUID:21396509; PMID:11481432
A;Accession: F95271
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-237 <KUR>
A;Cross-references: UNIPROT:Q930W6; UNIPARC:UPI000000CAF86; GB:AE006469; PIDN:AAK64736.1;
A;Experimental source: strain 1021, megaplasmid pSymA
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, R.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T. Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, J.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K. A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMA0148
A;Genome: plasmid

Query Match 83.8%; Score 31; DB 2; Length 237;
Best Local Similarity 57.1%; Pred. No. 1.8e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
:|||||
Db 23 VIAPPLP 29

RESULT 44
T42539
hypothetical protein - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C;Accession: T42539
R;Yoshioka, S.; Kato, K.; Nakai, K.; Okayama, H.; Nojima, H. DNA Res. 4, 363-369, 1997
A;Title: Identification of open reading frames in Schizosaccharomyces pombe cDNAs.
A;Reference number: Z17323; MUID:98162722; PMID:9501991
A;Accession: T42539
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-239 <YOS>
A;Cross-references: UNIPROT:P78826; UNIPARC:UPI000006B5DC; EMBL:D89174; NID:g1749555; PIDN:10-632-388-296
A;Experimental source: strain PR745

Query Match 83.8%; Score 31; DB 2; Length 239;

Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
:|||||
Db 17 IICPPVP 23

RESULT 45
B42768
homeotic protein goosecoid - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: B42768; D40856
R;Blum, M.; Gaunt, S.J.; Cho, K.W.Y.; Steinbeisser, H.; Blumberg, B.; Bittner, D.; De Robertis, E.M. Cell 69, 1097-1106, 1992
A;Title: Gastrulation in the mouse: the role of the homeobox gene goosecoid.
A;Reference number: A42768; MUID:92315328; PMID:1352187
A;Accession: B42768
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-243 <BLU>
A;Cross-references: UNIPARC:UPI00001745C0; GB:M63872
R;Blumberg, B.; Wright, C.V.E.; De Robertis, E.M.; Cho, K.W.Y. Science 253, 194-196, 1991
A;Title: Organizer-specific homeobox genes in Xenopus laevis embryos.
A;Reference number: A40856; MUID:91305940; PMID:1677215
A;Accession: D40856
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 148-207 <BL2>
A;Cross-references: UNIPARC:UPI00001745C1; GB:M63872
C;Superfamily: homeotic protein goosecoid; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;149-205/Domain: homeobox homology <HOX>

Query Match 83.8%; Score 31; DB 1; Length 243;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
:|||||
Db 114 VLMPPVP 120

RESULT 46
I51424
homeotic protein goosecoid [similarity] - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C;Accession: I51424
R;Blumberg, B.; Cho, K.W.Y.; Steinbeisser, H.; DeRobertis, E.M. Cell 67, 1111-1120, 1991
A;Title: Molecular nature of Spemann's organizer: The role of the Xenopus homeobox gene
A;Reference number: I51424; MUID:92103677; PMID:1684739
A;Accession: I51424
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-243 <BLU>
A;Cross-references: UNIPROT:P53546; UNIPARC:UPI000012BBA1; GB:M81481; NID:g214185; PIDN:10-632-388-296
C;Superfamily: homeotic protein goosecoid; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;149-205/Domain: homeobox homology <HOX>

Query Match 83.8%; Score 31; DB 2; Length 243;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
:|||||
Db 114 VLMPPVP 120

RESULT 47
AF2138
hypothetical protein all2661 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AF2138
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpō, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AF2138
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-326 <KUR>
A;Cross-references: UNIPROT:Q8YTQ2; UNIPARC:UPI00000CE4FB; GB:BA000019; PIDN:BA074360.1;
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all2661

Query Match 83.8%; Score 31; DB 2; Length 326;
Best Local Similarity 71.4%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
|:|||||
Db 189 ITAPPVP 195

RESULT 48
T51269
hypothetical protein T8M16_140 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
C;Accession: T51269
R;Benes, V.; Wurmbach, E.; Drzonek, H.; Ansorge, W.; Meves, H.W.; Rudd, S.; Lemcke, K.;
submitted to the Protein Sequence Database, August 2000
A;Reference number: Z25346
A;Accession: T51269
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-332 <BEN>
A;Cross-references: UNIPROT:Q9LES7; UNIPARC:UPI00000A48CC; EMBL:AL390921
A;Experimental source: cultivar Columbia; BAC clone T8M16
C;Genetics:
A;Map position: 3
A;Introns: 150/1
A;Note: T8M16_140

Query Match 83.8%; Score 31; DB 2; Length 332;
Best Local Similarity 71.4%; Pred. No. 2.6e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
|:|||||
Db 83 IVAPPLP 89

RESULT 49
A96835
gibberellin 3 beta-hydroxylase, 29683-28215 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: A96835
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: A96835
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-347 <STO>
A;Cross-references: UNIPROT:Q9C970; UNIPARC:UPI00000484C6; GB:AE005173; NID:g6751709; PI
C;Genetics:
A;Gene: F516.9
A;Map position: 1
C;Superfamily: 1-aminocyclopropane-1-carboxylate oxidase

Query Match 83.8%; Score 31; DB 2; Length 347;
Best Local Similarity 71.4%; Pred. No. 2.7e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
|:|||||
Db 252 VTAPPVP 258

RESULT 50
T51691
gibberellin 3 beta-hydroxylase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
C;Accession: T51691
R;Yamaguchi, S.; Smith, M.W.; Brown, R.G.; Kamiya, Y.; Sun, T.
Plant Cell 10, 2115-2126, 1998
A;Title: Phytochrome regulation and differential expression of gibberellin 3beta-hydroxy
A;Reference number: Z25425; MUID:99055200; PMID:9836749
A;Accession: T51691
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-347 <YAM>
A;Cross-references: UNIPROT:Q9ZT84; UNIPARC:UPI00000A0C66; EMBL:AF070937; PIDN:AAC83647.
C;Genetics:
A;Gene: GA4H
C;Superfamily: 1-aminocyclopropane-1-carboxylate oxidase

Query Match 83.8%; Score 31; DB 2; Length 347;
Best Local Similarity 71.4%; Pred. No. 2.7e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
|:|||||
Db 252 VTAPPVP 258

Search completed: April 6, 2006, 09:40:04
Job time : 24.7105 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2006, 09:23:24 ; Search time 85.1053 Seconds
(without alignments)
58.030 Million cell updates/sec

Title: US-10-632-388-296
Perfect score: 37
Sequence: 1 ILAPPVP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : UniProt 05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	100.0	221	Q5MZB3_SYPN6	Q5mzb3 synechococc
2	37	100.0	370	Q9LGV4_ORYSA	Q9lgv4 oryza sativ
3	37	100.0	1064	Q8SXB2_DROME	Q8sxb2 drosophila
4	37	100.0	1607	Q8IRS0_DROME	Q8irs0 drosophila
5	37	100.0	1767	Q24495_DROME	Q24495 drosophila
6	37	100.0	1767	Q9W4F5_DROME	Q9w4f5 drosophila
7	36	97.3	161	Q39063_ARATH	Q39063 arabidopsis
8	36	97.3	218	Q6ZN89_HUMAN	Q6zn89 homo sapien
9	36	97.3	461	Q7SBK3_NEUCR	Q7sbk3 neurospora
10	35	94.6	75	Q6Z8U5_ORYSA	Q6z8u5 oryza sativ
11	35	94.6	186	Q5D5E5_9RICK	Q5d5e5 wolbachia e
12	35	94.6	191	Q4UL63_RICFE	Q4ul63 rickettsia
13	35	94.6	191	Q73HK2_WOLPM	Q73hk2 wolbachia p
14	35	94.6	201	Q5PAP8_ANAMM	Q5pap8 anaplasma m
15	35	94.6	231	Q60W03_CAEBR	Q60w03 caenorhabdi
16	35	94.6	266	Q7D5X3_MYCTU	Q7d5x3 mycobacteri
17	35	94.6	266	Q05853_MYCTU	Q05853 mycobacteri
18	35	94.6	266	Q73UQ2_MYCPA	Q73uq2 mycobacteri
19	35	94.6	266	Q7TWZ9_MYCBO	Q7twz9 mycobacteri
20	35	94.6	477	Q5WMS2_ORYSA	Q5wms2 oryza sativ
21	35	94.6	662	Q7PYU0_ANOGA	Q7pyu0 anopheles g
22	35	94.6	842	Q55ZY4_CRYNE	Q55zy4 cryptococcu
23	35	94.6	842	Q5KPP2_CRYNE	Q5kp92 cryptococcu
24	35	94.6	990	Q6Y8G9_LEIME	Q6y8g9 leishmania
25	35	94.6	990	Q4QGY0_LEIMA	Q4qgy0 leishmania
26	34	91.9	189	PAAD_RICPR	Q9zd09 rickettsia
27	34	91.9	190	Q68WJ3_RICTY	Q68wj3 rickettsia
28	34	91.9	315	Q62I56_BURMA	Q62i56 burkholderi
29	34	91.9	341	Q5RI92_BRARE	Q5ri92 brachydanio
30	34	91.9	467	Q4SLJ8_TETNG	Q4slj8 tetraodon n
31	34	91.9	621	Q4IM73_GIBZE	Q4im73 gibberella

32	91.9	762	2	Q8T3T1_LYTVA	Q8t3t1 lytechinus
33	91.9	768	2	Q8C585_MOUSE	Q8c585 mus musculus
34	91.9	924	2	P73172_SYNY3	P73172 synechocyst
35	91.9	959	2	Q520P7_MAGGR	Q520p7 magnaporthe
36	91.9	1011	2	Q52FV3_MAGGR	Q52fv3 magnaporthe
37	91.9	1048	2	Q80TV1_MOUSE	Q80tv1 mus musculu
38	91.9	1291	2	Q6Y7W8_MOUSE	Q6y7w8 mus musculu
39	89.2	111	2	Q857Z6_9CAUD	Q857z6 mycobacteri
40	89.2	113	1	Y124_LISIN	Q925x3 listeria in
41	89.2	129	2	Q8DKV4_SYNEL	Q8dkv4 synechococc
42	89.2	147	2	Q9R6A1_9RHIZ	Q9r6a1 agrobacteri
43	89.2	164	2	Q9K3W9_STRCO	Q9k3w9 streptomyce
44	89.2	177	2	Q6Z9K3_ORYSA	Q6z9k3 oryza sativ
45	89.2	183	2	Q73TA7_MYCPA	Q73ta7 mycobacteri
46	89.2	205	1	HLH4_CAEL	P34555 caenorhabdi
47	89.2	226	2	P97177_RHOSH	P97177 rhodobacter
48	89.2	228	2	Q5FC67_CAEL	Q5fc67 caenorhabdi
49	89.2	235	2	Q52WY3_BRARE	Q52wy3 brachydanio
50	89.2	236	2	Q89E19_BRAJA	Q89e19 bradyrhizob
51	89.2	237	2	Q568D9_BRARE	Q568d9 brachydanio
52	89.2	258	2	Q9L1H7_STRCO	Q9l1h7 streptomyce
53	89.2	264	2	Q82C90_STRAW	Q82c90 streptomyce
54	89.2	281	2	Q82G67_STRAW	Q82g67 streptomyce
55	89.2	284	2	Q7XED3_ORYSA	Q7xed3 oryza sativ
56	89.2	315	2	Q63RF2_BURPS	Q63rf2 burkholderi
57	89.2	317	2	O52286_9RHIZ	O52286 agrobacteri
58	89.2	324	2	Q84YF5_SORBI	Q84yf5 sorghum bic
59	89.2	325	2	Q84YF3_SORBI	Q84yf3 sorghum bic
60	89.2	341	2	Q7R666_GIALA	Q7r666 giardia lam
61	89.2	370	2	Q8LGL0_SILDI	Q8lgl0 silene dioi
62	89.2	371	2	Q5K483_SILDI	Q5k483 silene dioi
63	89.2	375	2	Q4Q8Y4_LEIMA	Q4q8y4 leishmania
64	89.2	393	2	Q9EWE0_STRCO	Q9ewe0 streptomyce
65	89.2	418	2	Q4IBX5_GIBZE	Q4ibx5 gibberella
66	89.2	430	2	Q4HY43_GIBZE	Q4hy43 gibberella
67	89.2	444	2	Q965C9_GIALA	Q965c9 giardia lam
68	89.2	444	2	Q4TP09_9SPHN	Q4tp09 erythrobact
69	89.2	450	2	Q6MYW6_ASPFU	Q6myw6 aspergillus
70	89.2	452	2	Q7SCZ7_NEUCR	Q7scz7 neurospora
71	89.2	455	2	Q5BEB0_EMENI	Q5beb0 aspergillus
72	89.2	455	2	Q4WST6_ASPFU	Q4wst6 aspergillus
73	89.2	464	2	Q75HS1_ORYSA	Q75hs1 oryza sativ
74	89.2	465	2	O54176_STRCO	O54176 streptomyce
75	89.2	480	2	Q7RSN0_GIALA	Q7rsn0 giardia lam
76	89.2	483	2	Q5AZ77_EMENI	Q5az77 aspergillus
77	89.2	497	2	Q51HI0_MAGGR	Q51hi0 magnaporthe
78	89.2	506	2	Q6TV65_9POXV	Q6tv65 bovine papu
79	89.2	509	2	Q50LG1_ALTAL	Q50lg1 alternaria
80	89.2	516	2	Q51X61_MAGGR	Q51x61 magnaporthe
81	89.2	520	2	Q7F595_ORYSA	Q7f595 oryza sativ
82	89.2	520	2	Q9SLZ5_ORYSA	Q9slz5 oryza sativ
83	89.2	531	2	Q6C5V4_YARLI	Q6c5v4 yarrowia li
84	89.2	532	2	Q6TVJ6_9POXV	Q6tvj6 orf virus.
85	89.2	532	2	Q6TVX6_9POXV	Q6tvx6 orf virus.
86	89.2	535	2	Q6AYG6_RAT	Q6ayg6 rattus norv
87	89.2	538	2	P74506_SYNY3	P74506 synechocyst
88	89.2	540	2	Q5JLS2_ORYSA	Q5jls2 oryza sativ
89	89.2	543	2	Q8YY82_ANASP	Q8yy82 anabaena sp
90	89.2	551	2	Q46113_DROME	Q46113 drosophila
91	89.2	566	2	Q5WN70_CAEBR	Q5wn70 caenorhabdi
92	89.2	568	2	Q81ZZ0_STRAW	Q81zz0 streptomyce
93	89.2	572	2	Q4IJP9_GIBZE	Q4ijp9 gibberella
94	89.2	588	2	Q5B1Q1_EMENI	Q5blq1 aspergillus
95	89.2	592	2	Q7NCL0_GLOVI	Q7ncl0 gloebacter
96	89.2	597	2	Q6ZH89_ORYSA	Q6zh89 oryza sativ
97	89.2	601	2	Q4HA07_9DEIO	Q4ha07 deinococcus
98	89.2	604	2	Q4S2Y2_TETNG	Q4s2y2 tetraodon n
99	89.2	635	2	Q67UX0_ORYSA	Q67ux0 oryza sativ
100	89.2	637	2	Q6IVG5_EMENI	Q6ivg5 emericea
101	89.2	640	2	Q5ZJW0_CHICK	Q5zjw0 gallus gall
102	89.2	644	2	Q5NKM3_CRYNV	Q5nkm3 cryptococcu
103	89.2	647	2	Q5R9F2_PONPY	Q5r9f2 pongo pygma
104	89.2	648	2	Q7PXC5_ANOGA	Q7pxc5 anopheles g

105	33	89.2	687	2	Q59895	PODAN	Q59895	podospora a
106	33	89.2	766	2	Q52EX7	MAGGR	Q52ex7	magnaporthes
107	33	89.2	768	2	Q58LD2	9CAUD	Q58ld2	cyanophages
108	33	89.2	779	2	Q4Q991	LEIMA	Q4q991	leishmania
109	33	89.2	783	2	Q4WLT7	ASPFU	Q4wlt7	aspergillus
110	33	89.2	786	2	Q4Tl61	TETNG	Q4t161	tetraodon n
111	33	89.2	819	1	FTSK	CAUCR	Q9a262	caulobacter
112	33	89.2	830	2	Q7S967	NEUCR	Q7s967	neurospora
113	33	89.2	835	2	Q6K317	ORYSA	Q6k317	oryza sativ
114	33	89.2	853	2	Q59GQ3	HUMAN	Q59gq3	homo sapien
115	33	89.2	861	2	Q69JZ8	ORYSA	Q69jz8	oryza sativ
116	33	89.2	870	1	MILK1	MOUSE	Q8bgt6	mus musculu
117	33	89.2	905	2	Q658W4	HUMAN	Q658w4	homo sapien
118	33	89.2	920	2	Q4P159	USTMA	Q4p159	ustilago ma
119	33	89.2	962	2	Q7Z2X3	HUMAN	Q7z2x3	homo sapien
120	33	89.2	977	1	MCR	TUEGB	Q29131	tupaia glis
121	33	89.2	978	1	MCR	MOUSE	Q8vii8	mus musculu
122	33	89.2	980	1	RIN3	MOUSE	P59729	mus musculu
123	33	89.2	980	2	Q80WAO	MOUSE	Q80wa0	mus musculu
124	33	89.2	981	1	MCR	RAT	P22199	rattus norv
125	33	89.2	982	1	MCR	SAISC	Q9n0w8	saimiri sci
126	33	89.2	982	2	Q4Jm28	9PRIM	Q4jm28	saimiri bol
127	33	89.2	983	2	Q6XLI9	CALJA	Q6xli9	callithrix
128	33	89.2	984	1	MCR	HUMAN	P08235	homo sapien
129	33	89.2	987	2	Q6XLI8	CALJA	Q6xli8	callithrix
130	33	89.2	1092	2	Q4NN66	9DELT	Q4nn66	anaeromyxob
131	33	89.2	1094	2	Q5B395	EMENI	Q5b395	aspergillus
132	33	89.2	1214	2	Q73835	MYCPA	Q73835	mycobacteri
133	33	89.2	1252	2	Q4PC18	USTMA	Q4pc18	ustilago ma
134	33	89.2	1307	2	Q9V4J6	DROME	Q9v4j6	drosophila
135	33	89.2	1329	1	GP124	MOUSE	Q91zv8	mus musculu
136	33	89.2	1331	1	GP124	HUMAN	Q96pe1	homo sapien
137	33	89.2	1337	2	Q4Q3N2	LEIMA	Q4q3n2	leishmania
138	33	89.2	1367	2	Q20120	CABEL	Q20120	caenorhabdi
139	33	89.2	1413	2	Q582N5	9TRYYP	Q582n5	trypanosoma
140	33	89.2	1489	2	Q4PE83	USTMA	Q4pe83	ustilago ma
141	33	89.2	1725	2	Q62IF2	BURMA	Q62if2	burkholderi
142	33	89.2	1814	2	Q6WAY7	PEA	Q6way7	pisum sativ
143	33	89.2	1841	2	Q63VY6	BURPS	Q63vy6	burkholderi
144	33	89.2	2137	2	Q84BQ6	9PSED	Q84bq6	pseudomonas
145	33	89.2	2147	2	Q9L950	PSEPU	Q9l950	pseudomonas
146	33	89.2	2348	2	Q5Z1P8	NOCFA	Q5z1p8	nocardia fa
147	33	89.2	2518	2	Q8GBX7	POLCB	Q8gbx7	polyangium
148	33	89.2	6310	2	Q88PP2	PSEPK	Q88pp2	pseudomonas
149	33	89.2	7599	2	Q4IKE3	GIBZE	Q4ike3	gibberella
150	33	89.2	7962	2	Q10465	HUMAN	Q10465	homo sapien
151	32	86.5	51	2	Q865J2	LEMCA	Q865j2	lemur catta
152	32	86.5	57	2	Q865J3	LEMVA	Q865j3	lemur varie
153	32	86.5	57	2	Q6H3Y1	ORYSA	Q6h3y1	oryza sativ
154	32	86.5	68	2	Q5L1C0	MAGGR	Q5lic0	magnaporthes
155	32	86.5	101	2	Q8E5G8	STRA5	Q8e5g8	streptococc
156	32	86.5	101	2	Q7WQ05	BORBR	Q7wq05	bordetella
157	32	86.5	109	2	Q98MR9	RHILO	Q98mr9	rhizobium l
158	32	86.5	109	2	Q6ILY5	DROME	Q6ily5	drosophila
159	32	86.5	115	2	Q958K5	RANAU	Q958k5	rana aurora
160	32	86.5	121	2	Q69U83	ORYSA	Q69u83	oryza sativ
161	32	86.5	123	2	Q5N221	SYNP6	Q5n221	synechococc
162	32	86.5	125	2	Q60HI6	ONCKI	Q60hi6	oncorhynch
163	32	86.5	127	2	Q5NAP5	ORYSA	Q5nap5	oryza sativ
164	32	86.5	128	2	Q8N1G8	HUMAN	Q8nl98	homo sapien
165	32	86.5	157	2	Q60HI5	ONCTS	Q60hi5	oncorhynch
166	32	86.5	164	2	Q61UG9	DROME	Q61jg9	drosophila
167	32	86.5	165	2	Q5GTL0	WOLTR	Q5gtl0	wolbachia s
168	32	86.5	187	2	Q7P8V3	RICSI	Q7p8v3	rickettsia
169	32	86.5	191	2	Q92HH0	RICCN	Q92hh0	rickettsia
170	32	86.5	191	2	Q89DI5	BRAJA	Q89di5	bradyrhizob
171	32	86.5	194	2	Q86YU6	HUMAN	Q86yu6	homo sapien
172	32	86.5	200	2	Q89PF2	BRAJA	Q89pf2	bradyrhizob
173	32	86.5	206	2	Q9YCV5	ABRPE	Q9ycv5	aeropyrum p
174	32	86.5	210	2	Q6L5A6	ORYSA	Q6l5a6	oryza sativ
175	32	86.5	226	2	Y132	NEVOP	P24080	orgyia pseu
176	32	86.5	228	1	Q5YUT5	NOCFA	Q5yut5	nocardia fa
177	32	86.5	233	2				

178	32	86.5	297	2	Q8LA38	ARATH	Q8la38	arabidopsis
179	32	86.5	307	2	Q5S7Y6	PORYE	Q5s7y6	porphyra ye
180	32	86.5	310	2	Q527E0	MAGGR	Q527e0	magnaporthes
181	32	86.5	317	2	Q4QBD0	LEIMA	Q4qbd0	leishmania
182	32	86.5	324	2	Q90Z55	SCOMX	Q90z55	scophthalmu
183	32	86.5	335	2	Q5LV54	SILPO	Q5lv54	silicibacte
184	32	86.5	358	2	Q4KY19	ACASC	Q4ky19	acanthopagr
185	32	86.5	374	2	Q7S7V1	NEUCR	Q7s7v1	neurospora
186	32	86.5	402	2	Q5WQW1	CRYNE	Q5wqw1	cryptococcu
187	32	86.5	409	2	Q4H7W5	9DEIO	Q4h7w5	deinococcu
188	32	86.5	422	2	Q6A8M0	PROAC	Q6a8m0	propionibac
189	32	86.5	425	2	Q4NNI2	9DELT	Q4nni2	anaeromyxob
190	32	86.5	426	2	Q55PJ3	CRYNE	Q55pj3	cryptococcu
191	32	86.5	428	2	Q5KDW3	CRYNE	Q5kdw3	cryptococcu
192	32	86.5	428	2	Q4RPX5	TETNG	Q4rpx5	tetraodon n
193	32	86.5	430	2	Q8X0I6	NEUCR	Q8x0i6	neurospora
194	32	86.5	433	2	Q94IQ9	MAIZE	Q94iq9	zea mays (m
195	32	86.5	439	2	Q51KW7	MAGGR	Q51kw7	magnaporthes
196	32	86.5	460	2	Q9VM57	DROME	Q9vw57	drosophila
197	32	86.5	464	2	Q52CE0	MAGGR	Q52ce0	magnaporthes
198	32	86.5	469	2	Q4RKM8	TETNG	Q4rkm8	tetraodon n
199	32	86.5	471	2	Q9E1Z8	9ALPH	Q9e1z8	cercopithe
200	32	86.5	481	2	Q9SY17	ARATH	Q9sy17	arabidopsis
201	32	86.5	497	2	Q4T3X4	TETNG	Q4t3x4	tetraodon n
202	32	86.5	499	2	Q58DJ0	BOVIN	Q58dj0	bos taurus
203	32	86.5	510	2	Q9SP15	9LAMI	Q9sp15	asarina bar
204	32	86.5	518	2	Q5AZ81	EMENI	Q5az81	aspergillus
205	32	86.5	519	2	P93833	ARATH	P93833	arabidopsis
206	32	86.5	524	2	Q9VUU9	DROME	Q9vvu9	drosophila
207	32	86.5	530	2	Q7QFM7	ANOGA	Q7qfm7	anopheles g
208	32	86.5	557	2	Q50EX8	SPAAU	Q50ex8	sparus aura
209	32	86.5	563	2	Q4PC28	USTMA	Q4pc28	ustilago ma
210	32	86.5	567	2	Q6CAS9	YARLI	Q6cas9	yarrowia li
211	32	86.5	575	2	Q5NVA9	PONPY	Q5nva9	pongo pygma
212	32	86.5	583	2	Q50H40	SPAAU	Q50h40	sparus aura
213	32	86.5	590	2	Q68U44	SCOMX	Q68u44	scophthalmu
214	32	86.5	597	2	Q9C586	ARATH	Q9c586	arabidopsis
215	32	86.5	602	2	Q7T189	SILME	Q7t189	silurus mer
216	32	86.5	605	2	Q4R7G9	MACFA	Q4r7g9	macaca fasc
217	32	86.5	609	1	PPCK	STRCO	Q93jj15	streptomyce
218	32	86.5	611	2	Q8W420	ARATH	Q8w420	arabidopsis
219	32	86.5	613	2	Q9H9G8	HUMAN	Q9h9g8	homo sapien
220	32	86.5	623	2	Q6UAQ0	TETNG	Q6uaq0	tetraodon n
221	32	86.5	633	2	Q90Z56	SCOMX	Q90z56	scophthalmu
222	32	86.5	640	2	Q8AXA5	ACASC	Q8axa5	acanthopagr
223	32	86.5	642	2	Q23673	ARATH	Q23673	arabidopsis
224	32	86.5	647	2	Q71MN2	SPAAU	Q71mn2	sparus aura
225	32	86.5	657	2	Q60HI7	ONCMA	Q60hi7	oncorhynch
226	32	86.5	667	2	Q4X0B7	ASPFU	Q4x0b7	aspergillus
227	32	86.5	693	2	Q8WVW4	HUMAN	Q8wvw4	homo sapien
228	32	86.5	708	2	Q8V616	9VIRU	Q8v616	trichomonas
229	32	86.5	720	2	Q9CA22	ARATH	Q9ca22	arabidopsis
230	32	86.5	732	2	Q53T46	HUMAN	Q53t46	homo sapien
231	32	86.5	777	2	Q84VX4	ARATH	Q84vx4	arabidopsis
232	32	86.5	836	2	Q4SPR0	TETNG	Q4spr0	tetraodon n
233	32	86.5	864	2	Q4FYD4	LEIMA	Q4fyd4	leishmania
234	32	86.5	872	2	Q55RQ8	CRYNE	Q55rq8	cryptococcu
235	32	86.5	872	2	Q5KG63	CRYNE	Q5kg63	cryptococcu
236	32	86.5	879	2	Q6P553	MOUSE	Q6p553	mus musculu
237	32	86.5	885	2	Q4H2K1	CIOIN	Q4h2k1	ciona intes
238	32	86.5	894	2	Q5YT66	NOCFA	Q5yt66	nocardia fa
239	32	86.5	911	1	FTSK	SHEON	Q8eer3	shewanella
240	32	86.5	1048	2	Q9H0E3	HUMAN	Q9h0e3	homo sapien
241	32	86.5	1056	2	Q8BIH0	MOUSE	Q8bih0	mus musculu
242	32	86.5	1098	2	Q6NZP5	MOUSE	Q6nzp5	mus musculu
243	32	86.5	1100	2	Q9NV82	HUMAN	Q9nv82	homo sapien
244	32	86.5	1103	2	Q59FN7	HUMAN	Q59fn7	homo sapien
245	32	86.5	1148	2	Q7S9L4	NEUCR	Q7s9l4	neurospora
246	32	86.5	1202	2	Q5B6T7	EMENI	Q5b6t7	aspergillus
247	32	86.5	1298	2	Q7Z3I2	HUMAN	Q7z3i2	homo sapien
248	32	86.5	1299	2	Q6Y7W6	HUMAN	Q6y7w6	homo sapien
249	32	86.5	1299	2	Q7Z2Z8	HUMAN	Q7z2z8	homo sapien
250	32	86.5	1309	2	Q8GZA2	ARATH	Q8gza2	arabidopsis

251	32	86.5	1309	2	Q9LUG9_ARATH	Q9lug9 arabidopsis	31	83.8	237	2	Q930w6_RHIME	Q930w6 rhizobium m
252	32	86.5	1316	2	Q4IPP2_GIBZE	Q4ipp2 gibberella	31	83.8	239	1	YABF_SCHPO	YABF_SCHPO
253	32	86.5	1320	2	Q5A5B8_CANAL	Q5a5b8 candida alb	31	83.8	239	2	P78826_SCHPO	P78826 schizosacch
254	32	86.5	1329	2	Q75137_HUMAN	Q75137 homo sapien	31	83.8	243	1	GSCA_XENLA	GSCA_XENLA
255	32	86.5	1417	2	Q6C2D2_YARLI	Q6c2d2 yarrowia li	31	83.8	243	1	GSCB_XENLA	GSCB_XENLA
256	32	86.5	1448	2	Q4REE9_TETNG	Q4ree9 tetraodon n	31	83.8	243	2	Q68F73_XENLA	Q68f73 xenopus lae
257	32	86.5	1610	2	Q22173_CABEL	Q22173 caenorhabdi	31	83.8	243	2	Q4LDP9_XENTR	Q4ldp9 xenopus tro
258	32	86.5	1612	2	Q6BTM5_DEBHA	Q6btm5 debaryomyce	31	83.8	251	2	Q7NNL3_GLOVI	Q7nnl3 gloeobacter
259	32	86.5	1626	2	Q9RYY7_DEIRA	Q9ryy7 deinococcus	31	83.8	265	2	Q9X551_CORGL	Q9x551 corynebacte
260	32	86.5	1717	2	Q21875_CABEL	Q21875 caenorhabdi	31	83.8	270	2	Q75SV6_ONCMA	Q75sv6 oncorhynchu
261	32	86.5	1960	2	Q7R4L8_GIALA	Q7r4l8 giardia lam	31	83.8	274	2	Q73FZ3_WOLPM	Q73fz3 wolbachia p
262	32	86.5	2066	2	Q9Z3T8_PSESX	Q9z3t8 pseudomonas	31	83.8	275	2	Q5Z4H2_ORYSA	Q5z4h2 oryza sativ
263	32	86.5	2066	2	Q87W69_PSESM	Q87w69 pseudomonas	31	83.8	277	2	Q8RST7_ERWCH	Q8rst7 erwinia chr
264	32	86.5	2106	2	Q6C0Y6_YARLI	Q6c0y6 yarrowia li	31	83.8	291	2	Q51TQ6_MAGGR	Q51tq6 magnaporthe
265	32	86.5	2262	2	Q6WAY3_PEA	Q6way3 pisum sativ	31	83.8	294	2	Q8J243_9PLEO	Q8j243 phaeosphaer
266	32	86.5	3006	2	Q9Y4D8_HUMAN	Q9y4d8 homo sapien	31	83.8	295	2	Q8J244_9PLEO	Q8j244 phaeosphaer
267	32	86.5	3530	1	MYO15_HUMAN	Q9ukn7 homo sapien	31	83.8	298	2	Q4ITE0_AZOVI	Q4ite0 azotobacter
268	32	86.5	34350	2	Q8WZ42_HUMAN	Q8wz42 homo sapien	31	83.8	303	2	Q4K9B9_PSEF5	Q4k9b9 pseudomonas
269	31	83.8	52	2	Q6K2I5_ORYSA	Q6k2i5 oryza sativ	31	83.8	311	2	Q4IYM5_AZOVI	Q4iym5 azotobacter
270	31	83.8	110	2	Q5V4P1_HALMA	Q5v4p1 haloarcula	31	83.8	321	2	Q4SPI1_TETNG	Q4spi1 tetraodon n
271	31	83.8	123	2	Q4LXX5_9BURK	Q4lxx5 burkholderi	31	83.8	322	2	Q6Q1A6_9DELA	Q6q1a6 human t-lym
272	31	83.8	127	2	Q53WJ6_ORYSA	Q53wj6 oryza sativ	31	83.8	326	2	Q8YTQ2_ANASP	Q8ytq2 anabaena sp
273	31	83.8	133	2	Q59PD4_CANAL	Q59pd4 candida alb	31	83.8	328	2	Q4NJT6_9MICC	Q4njt6 arthrobacte
274	31	83.8	133	2	Q96VL8_CANAL	Q96vl8 candida alb	31	83.8	328	2	Q4M1A7_9BURK	Q4mla7 burkholderi
275	31	83.8	141	2	Q6QUJ6_9RHIZ	Q6quj6 hyphomicrob	31	83.8	329	2	Q9M5K1_EUPES	Q9m5k1 euphorbia e
276	31	83.8	142	2	Q6QUJ4_9RHIZ	Q6quj4 hyphomicrob	31	83.8	332	2	Q9LES7_ARATH	Q9les7 arabidopsis
277	31	83.8	142	2	Q6QUJ5_9RHIZ	Q6quj5 hyphomicrob	31	83.8	333	2	Q8KGB5_CHLTE	Q8kgb5 chlorobium
278	31	83.8	162	2	Q9A3B4_CAUCR	Q9a3b4 caulobacter	31	83.8	335	2	Q8B2W8_9PARA	Q8b2w8 measles vir
279	31	83.8	164	2	Q70R68_9HIV1	Q70r68 human immun	31	83.8	337	2	Q5EMQ9_9HIV1	Q5emq9 human immun
280	31	83.8	171	2	Q6EEU0_9HIV1	Q6eeu0 human immun	31	83.8	347	2	Q9C970_ARATH	Q9c970 arabidopsis
281	31	83.8	171	2	Q6EEU1_9HIV1	Q6eeu1 human immun	31	83.8	347	2	Q9ZT84_ARATH	Q9zt84 arabidopsis
282	31	83.8	178	1	DEF2_PSEPK	Q88ea7 pseudomonas	31	83.8	352	2	Q6PQQ9_ALLCE	Q6pqq9 allium cepa
283	31	83.8	179	1	DEF2_PSESM	Q886i1 pseudomonas	31	83.8	352	2	Q84TM0_ALLCE	Q84tm0 allium cepa
284	31	83.8	179	2	Q4ZPW1_PSESY	Q4zpw1 pseudomonas	31	83.8	361	2	Q57Z20_9TRYP	Q57z20 trypanosoma
285	31	83.8	179	2	Q4K7V5_PSEF5	Q4k7v5 pseudomonas	31	83.8	365	2	Q9WDB5_9HIV1	Q9wdb5 human immun
286	31	83.8	179	2	Q90CM9_9HIV1	Q90cm9 human immun	31	83.8	366	2	Q61852_CABEL	Q61852 caenorhabdi
287	31	83.8	185	2	Q6UDP6_9HIV1	Q6udp6 human immun	31	83.8	374	2	Q7ZP47_9HIV1	Q7zp47 human immun
288	31	83.8	189	2	Q5WST0_LEGPL	Q5wst0 legionella	31	83.8	379	2	Q4QHB1_LEIMA	Q4qhb1 leishmania
289	31	83.8	189	2	Q5X104_LEGPA	Q5x104 legionella	31	83.8	379	2	Q6N3A2_RHOPA	Q6n3a2 rhodopseudo
290	31	83.8	189	2	Q5ZRJ8_LEGPH	Q5zrj8 legionella	31	83.8	391	2	Q61V50_CABBR	Q61v50 caenorhabdi
291	31	83.8	193	2	Q70RC0_9HIV1	Q70rc0 human immun	31	83.8	392	2	Q4LQ85_9BURK	Q4lq85 burkholderi
292	31	83.8	193	2	Q9DVF3_9HIV1	Q9dvf3 human immun	31	83.8	396	2	Q87677_9PSEU	Q87677 amycolatops
293	31	83.8	194	2	Q7SKT2_9HIV1	Q7skt2 human immun	31	83.8	402	2	Q55972_Syny3	Q55972 synchocyst
294	31	83.8	195	2	Q7ZNJ2_9HIV1	Q7znj2 human immun	31	83.8	407	2	Q7NUV1_CHRVO	Q7nuv1 chromobacte
295	31	83.8	195	2	Q9DVA5_9HIV1	Q9dva5 human immun	31	83.8	407	2	Q4ZIU5_9HEPC	Q4zius hepatitis c
296	31	83.8	199	2	Q22262_CABEL	Q22262 caenorhabdi	31	83.8	410	1	ATIN_VZVD	P09265 varicella-z
297	31	83.8	199	2	Q7SKS5_9HIV1	Q7sk55 human immun	31	83.8	410	2	Q6QCF5_HHV3	Q6qcp5 human herpe
298	31	83.8	200	2	Q9DV90_9HIV1	Q9dv90 human immun	31	83.8	410	2	Q4JQW5_HHV3	Q4jqw5 human herpe
299	31	83.8	204	2	Q7ZNX3_9HIV1	Q7znx3 human immun	31	83.8	410	2	Q4JQY8_HHV3	Q4jqy8 human herpe
300	31	83.8	206	2	Q70R96_9HIV1	Q70r96 human immun	31	83.8	411	2	Q5VJ52_XENLA	Q5vj52 xenopus lae
301	31	83.8	208	2	Q8IP43_DROME	Q8ip43 drosophila	31	83.8	420	2	Q8DH06_SYNEL	Q8dh06 synchococc
302	31	83.8	210	2	Q92X27_RHIME	Q92x27 rhizobium m	31	83.8	421	2	Q6AZN0_XENLA	Q6azn0 xenopus lae
303	31	83.8	212	2	Q22263_CABEL	Q22263 caenorhabdi	31	83.8	422	2	Q4ID18_GIBZE	Q4id18 gibberella
304	31	83.8	213	2	Q8AD01_9HIV1	Q8ad01 human immun	31	83.8	423	2	Q5QIE2_CHICK	Q5qie2 gallus gall
305	31	83.8	214	2	Q5SMD6_THET8	Q5smd6 thermus the	31	83.8	425	2	Q6NZZ9_BRARE	Q6nzz9 brachydanio
306	31	83.8	214	2	Q72H80_THET2	Q72h80 thermus the	31	83.8	427	2	Q4LI48_9BURK	Q4li48 burkholderi
307	31	83.8	214	2	Q8XPW9_RALSO	Q8xpw9 ralstonia s	31	83.8	429	2	Q90CU2_9HIV1	Q90cu2 human immun
308	31	83.8	214	2	Q70QK9_9HIV1	Q70qk9 human immun	31	83.8	439	2	Q75HN4_ORYSA	Q75hn4 oryza sativ
309	31	83.8	217	2	Q4UZF3_XANCP	Q4uzf3 xanthomonas	31	83.8	445	2	Q8N597_HUMAN	Q8n597 homo sapien
310	31	83.8	217	2	Q8PD67_XANCP	Q8pd67 xanthomonas	31	83.8	445	2	Q9H8K7_HUMAN	Q9h8k7 homo sapien
311	31	83.8	217	2	Q9YXX4_9HIV1	Q9yxx4 human immun	31	83.8	457	2	Q4QUC9_9HIV1	Q4quc9 human immun
312	31	83.8	218	2	Q7ZBC7_9HIV1	Q7zbc7 human immun	31	83.8	459	2	Q4QTX1_9HIV1	Q4qtx1 human immun
313	31	83.8	219	2	Q7ZP23_9HIV1	Q7zp23 human immun	31	83.8	459	2	Q4QTW9_9HIV1	Q4qtw9 human immun
314	31	83.8	220	2	Q90R18_9HIV1	Q90r18 human immun	31	83.8	460	2	Q4QTY5_9HIV1	Q4qty5 human immun
315	31	83.8	220	2	Q9WLR8_9HIV1	Q9wlr8 human immun	31	83.8	462	2	Q5B0B8_EMENI	Q5b0b8 aspergillus
316	31	83.8	220	2	Q9YXX5_9HIV1	Q9yxx5 human immun	31	83.8	462	2	Q4QUF4_9HIV1	Q4quf4 human immun
317	31	83.8	220	2	Q9YXX5_9HIV1	Q9yxx5 human immun	31	83.8	462	2	Q4QUD1_9HIV1	Q4qud1 human immun
318	31	83.8	221	2	Q7ZMB5_9HIV1	Q7zmb5 human immun	31	83.8	462	2	Q4QUA3_9HIV1	Q4qua3 human immun
319	31	83.8	221	2	Q9QL77_9HIV1	Q9ql77 human immun	31	83.8	462	2	Q4QU99_9HIV1	Q4qu99 human immun
320	31	83.8	224	2	Q7ZMD7_9HIV1	Q7zmd7 human immun	31	83.8	462	2	Q4QU83_9HIV1	Q4qu83 human immun
321	31	83.8	224	2	Q90R25_9HIV1	Q90r25 human immun	31	83.8	462	2	Q4QU73_9HIV1	Q4qu73 human immun
322	31	83.8	225	2	Q9YXX9_9HIV1	Q9yxx9 human immun	31	83.8	462	2	Q4QU22_9HIV1	Q4qu22 human immun
323	31	83.8	228	2	Q90QY2_9HIV1	Q90qy2 human immun	31	83.8	462	2	Q4QU20_9HIV1	Q4qu20 human immun

397	31	83.8	462	2	Q4QU18_9HIV1	Q4qu18	human	immun
398	31	83.8	462	2	Q4QU12_9HIV1	Q4qu12	human	immun
399	31	83.8	462	2	Q4QU08_9HIV1	Q4qu08	human	immun
400	31	83.8	462	2	Q4QU06_9HIV1	Q4qu06	human	immun
401	31	83.8	462	2	Q4QT28_9HIV1	Q4qt28	human	immun
402	31	83.8	462	2	Q4QT24_9HIV1	Q4qt24	human	immun
403	31	83.8	462	2	Q4QT22_9HIV1	Q4qt22	human	immun
404	31	83.8	462	2	Q4QT20_9HIV1	Q4qt20	human	immun
405	31	83.8	462	2	Q4QTX9_9HIV1	Q4qtX9	human	immun
406	31	83.8	463	2	Q4QUA1_9HIV1	Q4qua1	human	immun
407	31	83.8	463	2	Q4QU97_9HIV1	Q4qu97	human	immun
408	31	83.8	463	2	Q4QU95_9HIV1	Q4qu95	human	immun
409	31	83.8	463	2	Q4QU69_9HIV1	Q4qu69	human	immun
410	31	83.8	463	2	Q4QU53_9HIV1	Q4qu53	human	immun
411	31	83.8	463	2	Q4QU47_9HIV1	Q4qu47	human	immun
412	31	83.8	463	2	Q4QU38_9HIV1	Q4qu38	human	immun
413	31	83.8	463	2	Q4QTW5_9HIV1	Q4qtW5	human	immun
414	31	83.8	469	2	Q58PX1_9HIV1	Q58px1	human	immun
415	31	83.8	470	2	Q9UVH4_SACPS	Q9uvh4	saccharomyc	
416	31	83.8	474	2	Q4QU93_9HIV1	Q4qu93	human	immun
417	31	83.8	474	2	Q4QU87_9HIV1	Q4qu87	human	immun
418	31	83.8	474	2	Q4QU75_9HIV1	Q4qu75	human	immun
419	31	83.8	474	2	Q4QU65_9HIV1	Q4qu65	human	immun
420	31	83.8	474	2	Q4QU63_9HIV1	Q4qu63	human	immun
421	31	83.8	474	2	Q4QU51_9HIV1	Q4qu51	human	immun
422	31	83.8	474	2	Q4QU49_9HIV1	Q4qu49	human	immun
423	31	83.8	474	2	Q4QU43_9HIV1	Q4qu43	human	immun
424	31	83.8	474	2	Q4QU41_9HIV1	Q4qu41	human	immun
425	31	83.8	474	2	Q4QU26_9HIV1	Q4qu26	human	immun
426	31	83.8	474	2	Q4QU24_9HIV1	Q4qu24	human	immun
427	31	83.8	476	2	Q4QTW1_9HIV1	Q4qtW1	human	immun
428	31	83.8	476	2	Q4QTV9_9HIV1	Q4qtV9	human	immun
429	31	83.8	476	2	Q4QTV7_9HIV1	Q4qtV7	human	immun
430	31	83.8	476	2	Q4QTV5_9HIV1	Q4qtV5	human	immun
431	31	83.8	476	2	Q4QTT9_9HIV1	Q4qtT9	human	immun
432	31	83.8	477	2	Q4QTV1_9HIV1	Q4qtV1	human	immun
433	31	83.8	481	2	Q4QU71_9HIV1	Q4qu71	human	immun
434	31	83.8	481	2	Q4QU77_9HIV1	Q4qu77	human	immun
435	31	83.8	481	2	Q4QTV3_9HIV1	Q4qtV3	human	immun
436	31	83.8	481	2	Q4QTV3_9HIV1	Q4qtu3	human	immun
437	31	83.8	481	2	Q4QTU1_9HIV1	Q4qtU1	human	immun
438	31	83.8	482	2	Q52GL7_MAGGR	Q52gl7	magnaporthes	
439	31	83.8	486	1	EGR4_HUMAN	Q052l5	homo sapien	
440	31	83.8	489	2	Q5CF52_CRYHO	Q5cfs2	cryptospori	
441	31	83.8	492	2	Q7ZV97_BRARE	Q7zv97	brachydanio	
442	31	83.8	492	2	Q98T96_BRARE	Q98t96	brachydanio	
443	31	83.8	492	2	Q85O91_9DELA	Q85091	simian t-ly	
444	31	83.8	496	2	Q5SHC3_THET8	Q5shc3	thermus the	
445	31	83.8	496	2	Q72HN1_THET2	Q72hn1	thermus the	
446	31	83.8	498	2	Q4I9A0_GIBZE	Q4i9a0	glibberella	
447	31	83.8	508	2	Q90CU3_9HIV1	Q90cu3	human	immun
448	31	83.8	510	2	Q55XN9_CRYNE	Q55xn9	cryptococcu	
449	31	83.8	510	2	Q5KM97_CRYNE	Q5km97	cryptococcu	
450	31	83.8	517	2	Q7VSA5_BORPE	Q7vsa5	bordetella	
451	31	83.8	522	2	Q5JLJ1_ORYSA	Q5jlJ1	oryza sativ	
452	31	83.8	524	2	Q9AJP0_STRGR	Q9ajP0	streptomyce	
453	31	83.8	524	2	Q7WC73_BORPA	Q7wc73	bordetella	
454	31	83.8	524	2	Q7WQ77_BORBR	Q7wq77	bordetella	
455	31	83.8	535	2	Q9IEF2_9HIV1	Q9ief2	human	immun
456	31	83.8	543	2	Q6FP63_CANGA	Q6fp63	candida gla	
457	31	83.8	553	2	Q9W4L5_DROME	Q9w4l5	drosophila	
458	31	83.8	557	2	Q5ZHN9_CHICK	Q5zhn9	gallus gall	
459	31	83.8	558	2	Q7SHL5_NEUCR	Q7shl5	neurospora	
460	31	83.8	565	2	Q8YB80_BRUME	Q8yb80	bruceella me	
461	31	83.8	565	2	Q8CEI3_MOUSE	Q8ceI3	mus musculu	
462	31	83.8	567	2	Q95RZ8_DROME	Q95rz8	drosophila	
463	31	83.8	573	2	Q89QM0_BRAJA	Q89qm0	bradyrhizob	
464	31	83.8	576	2	Q9IEF1_9HIV1	Q9ief1	human	immun
465	31	83.8	579	2	Q4S880_TETNG	Q4s880	tetradodon n	
466	31	83.8	591	2	Q50H41_ONCMY	Q50h41	oncorhynchu	
467	31	83.8	593	2	Q5I2A8_ONCMY	Q5i2a8	oncorhynchu	
468	31	83.8	593	2	Q90YP0_ONCKI	Q90yp0	oncorhynchu	
469	31	83.8	594	2	Q5J7P6_SALSA	Q5j7p6	salmo salar	

470	31	83.8	594	2	Q5MKL1_ONCMY	Q5mk11	oncorhynchu	
471	31	83.8	595	2	Q90W21_ONCMA	Q90w21	oncorhynchu	
472	31	83.8	595	2	Q90YN9_ONCKI	Q90yn9	oncorhynchu	
473	31	83.8	608	2	Q8FQH7_COREF	Q8fqh7	corynebacte	
474	31	83.8	619	2	Q23363_CAEBEL	Q23363	caenorhabdi	
475	31	83.8	620	2	Q7XVN3_ORYSA	Q7xvn3	oryza sativ	
476	31	83.8	629	2	Q8Y089_RALSO	Q8y089	raIstonia s	
477	31	83.8	629	2	Q6L631_ANGJA	Q6l631	anguilla ja	
478	31	83.8	644	2	Q6AF03_LEIXX	Q6af03	leifsonia x	
479	31	83.8	658	2	Q4WXC8_ASPFU	Q4wxc8	aspergillus	
480	31	83.8	660	2	Q9V5Z5_DROME	Q9v5z5	drosophila	
481	31	83.8	674	2	Q8U5E6_AGR75	Q8u5e6	agrobacteri	
482	31	83.8	680	2	Q729D8_DESVH	Q729d8	desulfovibr	
483	31	83.8	697	2	Q5AWZ1_EMENI	Q5awz1	aspergillus	
484	31	83.8	703	2	Q52CN4_MAGGR	Q52cn4	magnaporthes	
485	31	83.8	720	1	VP19_BPAPS	Q9t1s9	bacterioph	
486	31	83.8	722	2	Q7S563_NEUCR	Q7s563	neurospora	
487	31	83.8	724	2	Q8X0X3_NEUCR	Q8x0x3	neurospora	
488	31	83.8	728	2	Q8DJG9_SYNEL	Q8djg9	synechococc	
489	31	83.8	729	2	Q8S6N4_ORYSA	Q8s6n4	oryza sativ	
490	31	83.8	738	2	Q4PBG2_USTMA	Q4pbG2	ustilago ma	
491	31	83.8	748	2	Q55MI2_CRYNE	Q55mi2	cryptococcu	
492	31	83.8	760	2	Q5NB60_ORYSA	Q5nb60	oryza sativ	
493	31	83.8	826	2	Q96LM4_DROME	Q96lm4	drosophila	
494	31	83.8	847	2	Q69GT0_9HIV1	Q69gt0	human immun	
495	31	83.8	851	2	Q71AX6_9HIV1	Q71ax6	human immun	
496	31	83.8	854	2	Q9WSE4_9HIV1	Q9wse4	human immun	
497	31	83.8	854	2	Q9YV14_9HIV1	Q9yv14	human immun	
498	31	83.8	856	2	Q8JBY6_9HIV1	Q8jby6	human immun	
499	31	83.8	856	2	Q9WLH6_9HIV1	Q9wlh6	human immun	
500	31	83.8	858	2	Q4QWS7_9HIV1	Q4qws7	human immun	
501	31	83.8	859	2	Q8JBW8_9HIV1	Q8jbw8	human immun	
502	31	83.8	860	2	Q93091_9HIV1	Q93091	human immun	
503	31	83.8	863	2	Q6UPP6_9HIV1	Q6ufp6	human immun	
504	31	83.8	864	2	Q5MYI2_SIVCZ	Q5myi2	chimpanzee	
505	31	83.8	864	2	Q6UEP8_9HIV1	Q6uep8	human immun	
506	31	83.8	864	2	Q998H1_9HIV1	Q998h1	human immun	
507	31	83.8	865	2	Q9IV12_9HIV1	Q9iv12	human immun	
508	31	83.8	866	2	Q90D24_9HIV1	Q90d24	human immun	
509	31	83.8	876	2	Q79670_9HIV1	Q79670	human immun	
510	31	83.8	877	1	WEE1_SCHPO	P07527	schizosacch	
511	31	83.8	878	2	Q8JBD5_9HIV1	Q8jbd5	human immun	
512	31	83.8	880	2	Q8PPS5_XANAC	Q8pps5	xanthomonas	
513	31	83.8	900	2	Q8X0N9_NEUCR	Q8x0n9	neurospora	
514	31	83.8	919	2	Q51QS1_MAGGR	Q5lqs1	magnaporthes	
515	31	83.8	927	2	Q881W9_PSESM	Q881w9	pseudomonas	
516	31	83.8	937	2	Q5GVJ8_XANOR	Q5gvj8	xanthomonas	
517	31	83.8	958	2	Q77274_DROME	O77274	drosophila	
518	31	83.8	968	2	Q6ZIW1_ORYSA	Q6ziw1	oryza sativ	
519	31	83.8	978	2	Q9W4J0_DROME	Q9w4j0	drosophila	
520	31	83.8	993	2	Q73US5_MYCPA	Q73us5	mycobacteri	
521	31	83.8	1016	2	Q6ZTGH_HUMAN	Q6ztgh	homo sapien	
522	31	83.8	1016	2	Q9W4I9_DROME	Q9w4i9	drosophila	
523	31	83.8	1037	2	Q5CWR7_CRYPV	Q5cwr7	cryptospori	
524	31	83.8	1037	2	Q7YYN7_CRYPV	Q7yyN7	cryptospori	
525	31	83.8	1038	2	Q875P1_SACKL	Q875p1	saccharomyc	
526	31	83.8	1052	1	BUB1B_MOUSE	Q9z1s0	mus musculu	
527	31	83.8	1052	2	Q8K2E5_MOUSE	Q8k2e5	mus musculu	
528	31	83.8	1085	1	BUB1_HUMAN	Q43683	homo sapien	
529	31	83.8	1085	2	Q53QE4_HUMAN	Q53qe4	homo sapien	
530	31	83.8	1117	2	Q5KN38_CRYNE	Q5kn38	cryptococcu	
531	31	83.8	1117	2	Q55YR7_CRYNE	Q55yr7	cryptococcu	
532	31	83.8	1137	2	Q4N143_THEPA	Q4n143	thelateria p	
533	31	83.8	1189	2	Q4S1J4_TETNG	Q4s1j4	tetradodon n	
534	31	83.8	1201	2	Q51PJ1_MAGGR	Q51pj1	magnaporthes	
535	31	83.8	1228	2	Q5AGD7_CANAL	Q5agd7	candida alb	
536	31	83.8	1230	1	STE20_CANAL	Q922i2	candida alb	
537	31	83.8	1230	2	O1343I_CANAL	O1343i	candida alb	
538	31	83.8	1239	1	M4K4_HUMAN	Q958l9	homo sapien	
539	31	83.8	1249	2	Q4P3M6_USTMA	Q4p3m6	ustilago ma	
540	31	83.8	1260	2	Q53TX8_HUMAN	Q53tx8	homo sapien	
541	31	83.8	1262	2	Q9N3T9_CAEBEL	Q9n3t9	caenorhabdi	
542	31	83.8	1302	2	Q9W4I8_DROME	Q9w4i8	drosophila	

543	31	83.8	1306	2	077273	DROME	077273	drosophila	616	30	81.1	203	2	Q57E20	BRUAB	Q57E20	brucella ab
544	31	83.8	1322	2	Q5BI31	_DROME	Q5bi31	drosophila	617	30	81.1	203	2	Q8G1J2	_BRUSU	Q8glj2	brucella su
545	31	83.8	1325	2	Q871Q0	_NEUCR	Q871q0	neurospora	618	30	81.1	203	2	Q8YGD0	_BRUME	Q8ygd0	brucella me
546	31	83.8	1506	2	Q7PQL0	_ANOGA	Q7pqL0	anopheles g	619	30	81.1	203	2	Q4JLU1	_GOBNI	Q4jlu1	gobius nige
547	31	83.8	1530	1	RPC1	_TRYBB	P08968	trypanosoma	620	30	81.1	206	2	Q5H465	_XANOR	Q5h465	xanthomonas
548	31	83.8	1547	1	RRPO	_PMV	P20951	papaya mosa	621	30	81.1	207	2	Q5NEU6	_MOUSE	Q5nbu6	mus musculus
549	31	83.8	1567	2	Q6CB83	_YARLI	Q6cb83	yarrowia li	622	30	81.1	209	2	Q98B47	_RHILO	Q98b47	rhizobium l
550	31	83.8	1601	2	Q54L90	_DICDI	Q54l90	dictyosteli	623	30	81.1	212	2	Q9A7C0	_CAUCR	Q9a7c0	caulobacter
551	31	83.8	1611	2	Q42854	_SCHPO	Q42854	schizosacch	624	30	81.1	217	2	Q7X065	_9BACT	Q7x065	uncultured
552	31	83.8	1673	2	Q9VQI9	_DROME	Q9vqi9	drosophila	625	30	81.1	218	2	Q9MBH9	_9VIRU	Q9mbh9	bacteriopha
553	31	83.8	1687	2	Q4P9Z7	_USTMA	Q4p9z7	ustilago ma	626	30	81.1	218	2	Q8LMV5	_ORYSA	Q8lmv5	oryza sativ
554	31	83.8	1701	2	Q5Z1T0	_NOCPA	Q5z1t0	nocardia fa	627	30	81.1	220	2	Q88461	_9HERP	Q88461	stealth vir
555	31	83.8	1861	2	Q4WRV0	_ASPTU	Q4wrv0	aspergillus	628	30	81.1	221	2	Q50523	_METTF	Q50523	methanobact
556	31	83.8	2368	2	Q4S4K5	_TETNG	Q4s4k5	tetraodon n	629	30	81.1	227	2	Q6K4U0	_ORYSA	Q6k4u0	oryza sativ
557	31	83.8	2379	2	Q6U9W8	_PRRSV	Q6u9w8	porcine rep	630	30	81.1	227	2	Q54078	_RHOSH	Q54078	rhodobacter
558	31	83.8	2606	2	Q36414	_9GAMA	Q36414	alcelaphine	631	30	81.1	229	2	Q6JN61	_XANCA	Q6jn61	xanthomonas
559	31	83.8	3010	2	Q9QIX3	_9HEPC	Q9qix3	hepatitis c	632	30	81.1	230	2	Q5P6D9	_AZOSE	Q5p6d9	azoarcus sp
560	31	83.8	3010	2	Q9QIX4	_9HEPC	Q9qix4	hepatitis c	633	30	81.1	231	2	Q9XTR4	_CAEEL	Q9xtr4	caenorhabdi
561	31	83.8	4881	2	Q9S0R3	_STRAW	Q9s0r3	streptomyce	634	30	81.1	238	2	Q4I9Y6	_GIBZE	Q4i9y6	gibberella
562	31	83.8	6675	2	Q4KCD8	_PSEF5	Q4kcd8	pseudomonas	635	30	81.1	240	2	Q7EYB9	_ORYSA	Q7eyb9	oryza sativ
563	30	81.1	30	2	Q4YG77	_PLABE	Q4yg77	plasmodium	636	30	81.1	241	2	Q7PRT7	_ANOGA	Q7prt7	anopheles g
564	30	81.1	55	2	Q607C9	_METCA	Q607c9	methylococc	637	30	81.1	248	1	TPIS	_DEBHA	Q6bmb8	debaryomyce
565	30	81.1	73	2	Q9J2H5	_9GAMA	Q9j2h5	rhesus monk	638	30	81.1	248	2	Q5N4E5	_SYNP6	Q5n4e5	synecococc
566	30	81.1	75	2	Q93EZ3	_SHIFL	Q93ez3	shigella fl	639	30	81.1	250	2	Q9RSP1	_DEIRA	Q9rsp1	deinococcus
567	30	81.1	78	2	Q28373	_HORSE	Q28373	equus cabal	640	30	81.1	251	2	Q4R9D7	_MACFA	Q4r9d7	macaca fasc
568	30	81.1	84	2	Q4RZ95	_TETNG	Q4rz95	tetraodon n	641	30	81.1	254	2	Q6TZF4	_9HEPC	Q6tzf4	hepatitis c
569	30	81.1	85	2	Q5Z8E4	_ORYSA	Q5z8e4	oryza sativ	642	30	81.1	254	2	Q6U0C3	_9HEPC	Q6u0c3	hepatitis c
570	30	81.1	87	2	Q5JR97	_HUMAN	Q5jr97	homo sapien	643	30	81.1	254	2	Q6U4K1	_9HEPC	Q6u4k1	hepatitis c
571	30	81.1	102	2	Q4M0I3	_9BURK	Q4m0i3	burkholderi	644	30	81.1	254	2	Q6U4K2	_9HEPC	Q6u4k2	hepatitis c
572	30	81.1	108	2	Q68D33	_HUMAN	Q68d33	homo sapien	645	30	81.1	254	2	Q6U4K3	_9HEPC	Q6u4k3	hepatitis c
573	30	81.1	109	2	Q8WYT7	_HUMAN	Q8wyt7	homo sapien	646	30	81.1	254	2	Q6U4K4	_9HEPC	Q6u4k4	hepatitis c
574	30	81.1	115	2	Q9P7Y2	_SCHPO	Q9p7y2	schizosacch	647	30	81.1	254	2	Q6U4K5	_9HEPC	Q6u4k5	hepatitis c
575	30	81.1	121	2	Q66VC9	_LEIDO	Q66vc9	leishmania	648	30	81.1	254	2	Q6U4K6	_9HEPC	Q6u4k6	hepatitis c
576	30	81.1	121	2	Q8C3V1	_MOUSE	Q8c3v1	mus musculus	649	30	81.1	254	2	Q6U4K7	_9HEPC	Q6u4k7	hepatitis c
577	30	81.1	123	2	Q65178	_ASF	Q65178	african swi	650	30	81.1	254	2	Q6U4K8	_9HEPC	Q6u4k8	hepatitis c
578	30	81.1	128	2	Q8N8A0	_HUMAN	Q8n8a0	homo sapien	651	30	81.1	254	2	Q6U4K9	_9HEPC	Q6u4k9	hepatitis c
579	30	81.1	129	2	Q4UK21	_RICPE	Q4uk21	rickettsia	652	30	81.1	254	2	Q6U4L0	_9HEPC	Q6u4l0	hepatitis c
580	30	81.1	131	2	Q64DU2	_9ARCH	Q64du2	uncultured	653	30	81.1	254	2	Q6U4L1	_9HEPC	Q6u4l1	hepatitis c
581	30	81.1	132	2	Q5N5H8	_SYNP6	Q5n5h8	synecococc	654	30	81.1	254	2	Q6U4L3	_9HEPC	Q6u4l3	hepatitis c
582	30	81.1	134	2	Q5P3P8	_AZOSE	Q5p3p8	azoarcus sp	655	30	81.1	254	2	Q6U4L4	_9HEPC	Q6u4l4	hepatitis c
583	30	81.1	136	2	Q9Y9X1	_AERPE	Q9y9x1	aeropyrum p	656	30	81.1	254	2	Q6U4L5	_9HEPC	Q6u4l5	hepatitis c
584	30	81.1	139	2	Q4TG84	_TETNG	Q4tg84	tetraodon n	657	30	81.1	254	2	Q6U4L6	_9HEPC	Q6u4l6	hepatitis c
585	30	81.1	148	2	Q5CDG1	_CRYHO	Q5cdg1	cryptospori	658	30	81.1	254	2	Q6U4L7	_9HEPC	Q6u4l7	hepatitis c
586	30	81.1	149	2	Q5V1Z0	_HALMA	Q5v1z0	haloarcula	659	30	81.1	254	2	Q6U4L8	_9HEPC	Q6u4l8	hepatitis c
587	30	81.1	152	2	Q55YB3	_CRYNE	Q55yb3	cryptococcu	660	30	81.1	254	2	Q6U4L9	_9HEPC	Q6u4l9	hepatitis c
588	30	81.1	152	2	Q5KLM7	_CRYNE	Q5klm7	cryptococcu	661	30	81.1	254	2	Q6U4M0	_9HEPC	Q6u4m0	hepatitis c
589	30	81.1	152	2	Q7TFM0	_RHCM6	Q7tfm0	rhesus cyto	662	30	81.1	254	2	Q6U4M1	_9HEPC	Q6u4m1	hepatitis c
590	30	81.1	153	2	Q7F0Q3	_ORYSA	Q7f0q3	oryza sativ	663	30	81.1	254	2	Q6U4M3	_9HEPC	Q6u4m3	hepatitis c
591	30	81.1	156	2	Q9W8S9	_9BETA	Q9w8s9	human herpe	664	30	81.1	254	2	Q6U4M4	_9HEPC	Q6u4m4	hepatitis c
592	30	81.1	157	2	Q84XQ6	_BRARP	Q84xq6	brassica ra	665	30	81.1	254	2	Q6U4M5	_9HEPC	Q6u4m5	hepatitis c
593	30	81.1	160	2	Q4TJ93	_TETNG	Q4tj93	tetraodon n	666	30	81.1	254	2	Q6U4M6	_9HEPC	Q6u4m6	hepatitis c
594	30	81.1	163	2	Q7TWV7	_MYCBO	Q7twv7	mycobacteri	667	30	81.1	254	2	Q6U4M7	_9HEPC	Q6u4m7	hepatitis c
595	30	81.1	163	2	Q53351	_MYCTU	Q53351	mycobacteri	668	30	81.1	254	2	Q6U4M8	_9HEPC	Q6u4m8	hepatitis c
596	30	81.1	166	2	Q5IKZ7	_9HIV1	Q5ikz7	human immun	669	30	81.1	254	2	Q6U4M9	_9HEPC	Q6u4m9	hepatitis c
597	30	81.1	168	2	Q7EZX8	_ORYSA	Q7ezx8	oryza sativ	670	30	81.1	254	2	Q6U4N0	_9HEPC	Q6u4n0	hepatitis c
598	30	81.1	173	2	Q89Y86	_BRAJA	Q89y86	bradyrhizob	671	30	81.1	254	2	Q6U4N1	_9HEPC	Q6u4n1	hepatitis c
599	30	81.1	174	2	Q6EPY1	_ORYSA	Q6epy1	oryza sativ	672	30	81.1	254	2	Q6U4N2	_9HEPC	Q6u4n2	hepatitis c
600	30	81.1	175	2	Q5Z545	_ORYSA	Q5z545	oryza sativ	673	30	81.1	254	2	Q6U4N3	_9HEPC	Q6u4n3	hepatitis c
601	30	81.1	179	2	Q27803	_METHH	Q27803	methanobact	674	30	81.1	254	2	Q6U4N4	_9HEPC	Q6u4n4	hepatitis c
602	30	81.1	181	2	Q4TSG3	_9SPHN	Q4tsG3	erythrobaet	675	30	81.1	254	2	Q6U4N5	_9HEPC	Q6u4n5	hepatitis c
603	30	81.1	188	2	Q5TUP8	_ANOGA	Q5tup8	anopheles g	676	30	81.1	254	2	Q6U4N6	_9HEPC	Q6u4n6	hepatitis c
604	30	81.1	190	2	Q6D2P4	_BRWCT	Q6d2p4	erwinia car	677	30	81.1	254	2	Q6U4N7	_9HEPC	Q6u4n7	hepatitis c
605	30	81.1	191	2	Q6U2T2	_9HEPC	Q6u2t2	hepatitis c	678	30	81.1	254	2	Q6U4N8	_9HEPC	Q6u4n8	hepatitis c
606	30	81.1	191	2	Q6U2T8	_9HEPC	Q6u2t8	hepatitis c	679	30	81.1	254	2	Q6U4N9	_9HEPC	Q6u4n9	hepatitis c
607	30	81.1	194	2	Q9A6M7	_CAUCR	Q9a6m7	caulobacter	680	30	81.1	254	2	Q6U4P0	_9HEPC	Q6u4p0	hepatitis c
608	30	81.1	197	2	Q92XP7	_RHIME	Q92xp7	rhizobium m	681	30	81.1	254	2	Q6U4P1	_9HEPC	Q6u4p1	hepatitis c
609	30	81.1	198	2	Q96SJ7	_HUMAN	Q96sj7	homo sapien	682	30	81.1	254	2	Q6U4P2	_9HEPC	Q6u4p2	hepatitis c
610	30	81.1	199	2	Q7ZP17	_9HIV1	Q7zpl7	human immun	683	30	81.1	254	2	Q6U4P3	_9HEPC	Q6u4p3	hepatitis c
611	30	81.1	200	2	Q8GSQ9	_WHEAT	Q8gsq9	trititicum ae	684	30	81.1	254	2	Q6U4P5	_9HEPC	Q6u4p5	hepatitis c
612	30	81.1	201	2	Q57P18	_SALCH	Q57p18	salmonella	685	30	81.1	254	2	Q6U4P6	_9HEPC	Q6u4p6	hepatitis c
613	30	81.1	202	2	Q6K3R4	_ORYSA	Q6k3r4	oryza sativ	686	30	81.1	254	2	Q6U4P7	_9HEPC	Q6u4p7	hepatitis c
614	30	81.1	202	2	Q7V7H9	_PROMM	Q7v7h9	prochloroco	687	30	81.1	254	2	Q6U4Q4	_9HEPC	Q6u4q4	hepatitis c
615	30	81.1	202	2	Q70QX2	_9HIV1	Q70qx2	human immun	688	30	81.1	254	2	Q6U4R2	_9HEPC	Q6u4r2	hepatitis c

30	81.1	203	2	Q57E20	BRUAB	Q57e20	brucella ab
30	81.1	203	2	Q8G1J2	BRUSU	Q8glj2	brucella su
30	81.1	203	2	Q8YGD0	BRUME	Q8ygd0	brucella me
30	81.1	203	2	Q4JLU1	GOBNI	Q4jlu1	gobius nige
30	81.1	206	2	Q5H465	XANOR	Q5h465	xanthomonas
30	81.1	207	2	Q5NBU6	MOUSE	Q5nbu6	mus musculus
30	81.1	209	2	Q98B47	RHILO	Q98b47	rhizobium l
30	81.1	212	2	Q9A7C0	CAUCR	Q9a7c0	caulobacter
30	81.1	217	2	Q7X065	9BACT	Q7x065	uncultured
30	81.1	218	2	Q9MBH9	9VIRU	Q9mbh9	bacterioph
30	81.1	218	2	Q8LMV5	ORYSA	Q8lmv5	oryza sativ
30	81.1	220	2	Q88461	9HERP	Q88461	stealth vir
30	81.1	221	2	Q50523	METTF	Q50523	methanobact
30	81.1	227	2	Q6K4U0	ORYSA	Q6k4u0	oryza sativ
30	81.1	227	2	Q54078	RHOSH	Q54078	rhodobacter
30	81.1	229	2	Q6JN61	XANCA	Q6jn61	xanthomonas
30	81.1	230	2	Q5P6D9	AZOSE	Q5p6d9	azoarcus sp
30	81.1	231	2	Q9XTR4	CAEHL	Q9xtr4	caenorhabdi
30	81.1	238	2	Q4I9Y6	GIBZE	Q4i9y6	gibberella
30	81.1	240	2	Q7EYB9	ORYSA	Q7eyb9	oryza sativ
30	81.1	241	2	Q7PRT7	ANOGA	Q7prt7	anopheles g
30	81.1	248	1	TPIS	DEHA	Q6bmb8	debaryomyce
30	81.1	248	2	Q5N4E5	SYNP6	Q5n4e5	synchococc
30	81.1	250	2	Q9RSPI	DEIRA	Q9rsp1	deinococcus
30	81.1	251	2	Q4R9D7	MACFA	Q4r9d7	macaca fasc
30	81.1	254	2	Q6TZF4	9HEPC	Q6tzf4	hepatitis c
30	81.1	254	2	Q6U0C3	9HEPC	Q6u0c3	hepatitis c
30	81.1	254	2	Q6U4K1	9HEPC	Q6u4k1	hepatitis c
30	81.1	254	2	Q6U4K2	9HEPC	Q6u4k2	hepatitis c
30	81.1	254	2	Q6U4K3	9HEPC	Q6u4k3	hepatitis c
30	81.1	254	2	Q6U4K4	9HEPC	Q6u4k4	hepatitis c
30	81.1	254	2	Q6U4K5	9HEPC	Q6u4k5	hepatitis c
30	81.1	254	2	Q6U4K6	9HEPC	Q6u4k6	hepatitis c
30	81.1	254	2	Q6U4K7	9HEPC	Q6u4k7	hepatitis c
30	81.1	254	2	Q6U4K8	9HEPC	Q6u4k8	hepatitis c
30	81.1	254	2	Q6U4K9	9HEPC	Q6u4k9	hepatitis c
30	81.1	254	2	Q6U4L0	9HEPC	Q6u4l0	hepatitis c
30	81.1	254	2	Q6U4L1	9HEPC	Q6u4l1	hepatitis c
30	81.1	254	2	Q6U4L3	9HEPC	Q6u4l3	hepatitis c
30	81.1	254	2	Q6U4L4	9HEPC	Q6u4l4	hepatitis c
30	81.1	254	2	Q6U4L5	9HEPC	Q6u4l5	hepatitis c
30	81.1	254	2	Q6U4L6	9HEPC	Q6u4l6	hepatitis c
30	81.1	254	2	Q6U4L7	9HEPC	Q6u4l7	hepatitis c
30	81.1	254	2	Q6U4L8	9HEPC	Q6u4l8	hepatitis c
30	81.1	254	2	Q6U4L9	9HEPC	Q6u4l9	hepatitis c
30	81.1	254	2	Q6U4M0	9HEPC	Q6u4m0	hepatitis c
30	81.1	254	2	Q6U4M1	9HEPC	Q6u4m1	hepatitis c
30	81.1	254	2	Q6U4M3	9HEPC	Q6u4m3	hepatitis c
30	81.1	254	2	Q6U4M4	9HEPC	Q6u4m4	hepatitis c
30	81.1	254	2	Q6U4M5	9HEPC	Q6u4m5	hepatitis c
30	81.1	254	2	Q6U4M6	9HEPC	Q6u4m6	hepatitis c
30	81.1	254	2	Q6U4M7	9HEPC	Q6u4m7	hepatitis c
30	81.1	254	2	Q6U4M8	9HEPC	Q6u4m8	hepatitis c
30	81.1	254	2	Q6U4M9	9HEPC	Q6u4m9	hepatitis c
30	81.1	254	2	Q6U4N0	9HEPC	Q6u4n0	hepatitis c
30	81.1	254	2	Q6U4N1	9HEPC	Q6u4n1	hepatitis c
30	81.1	254	2	Q6U4N2	9HEPC	Q6u4n2	hepatitis c
30	81.1	254	2	Q6U4N3	9HEPC	Q6u4n3	hepatitis c
30	81.1	254	2	Q6U4N4	9HEPC	Q6u4n4	hepatitis c
30	81.1	254	2	Q6U4N5	9HEPC	Q6u4n5	hepatitis c
30	81.1	254	2	Q6U4N6	9HEPC	Q6u4n6	hepatitis c
30	81.1	254	2	Q6U4N7	9HEPC	Q6u4n7	hepatitis c
30	81.1	254	2	Q6U4N8	9HEPC	Q6u4n8	hepatitis c
30	81.1	254	2	Q6U4N9	9HEPC	Q6u4n9	hepatitis c
30	81.1	254	2	Q6U4P0	9HEPC	Q6u4p0	hepatitis c
30	81.1	254	2	Q6U4P1	9HEPC	Q6u4p1	hepatitis c
30	81.1	254	2	Q6U4P2	9HEPC	Q6u4p2	hepatitis c
30	81.1	254	2	Q6U4P3	9HEPC	Q6u4p3	hepatitis c
30	81.1	254	2	Q6U4P5	9HEPC	Q6u4p5	hepatitis c
30	81.1	254	2	Q6U4P6	9HEPC	Q6u4p6	hepatitis c
30	81.1	254	2	Q6U4P7	9HEPC	Q6u4p7	hepatitis c
30	81.1	254	2	Q6U4Q4	9HEPC	Q6u4q4	hepatitis c
30	81.1	254	2	Q6U4R2	9HEPC	Q6u4r2	hepatitis c

689	30	81.1	254	2	Q6U594_9HEPC	Q6u594	hepatitis c
690	30	81.1	254	2	Q6U5A3_9HEPC	Q6u5a3	hepatitis c
691	30	81.1	254	2	Q6U5A4_9HEPC	Q6u5a4	hepatitis c
692	30	81.1	254	2	Q6U5A5_9HEPC	Q6u5a5	hepatitis c
693	30	81.1	254	2	Q6U5A7_9HEPC	Q6u5a7	hepatitis c
694	30	81.1	254	2	Q6U5B0_9HEPC	Q6u5b0	hepatitis c
695	30	81.1	254	2	Q6U5B2_9HEPC	Q6u5b2	hepatitis c
696	30	81.1	254	2	Q6U5B4_9HEPC	Q6u5b4	hepatitis c
697	30	81.1	254	2	Q6U5B5_9HEPC	Q6u5b5	hepatitis c
698	30	81.1	254	2	Q6U5B6_9HEPC	Q6u5b6	hepatitis c
699	30	81.1	254	2	Q6U5B7_9HEPC	Q6u5b7	hepatitis c
700	30	81.1	254	2	Q6U5B8_9HEPC	Q6u5b8	hepatitis c
701	30	81.1	254	2	Q6U5C0_9HEPC	Q6u5c0	hepatitis c
702	30	81.1	254	2	Q6U5C4_9HEPC	Q6u5c4	hepatitis c
703	30	81.1	254	2	Q6U5C7_9HEPC	Q6u5c7	hepatitis c
704	30	81.1	254	2	Q6U5D3_9HEPC	Q6u5d3	hepatitis c
705	30	81.1	254	2	Q6U5D4_9HEPC	Q6u5d4	hepatitis c
706	30	81.1	254	2	Q6U5D5_9HEPC	Q6u5d5	hepatitis c
707	30	81.1	254	2	Q6U5D5_9HEPC	Q6u5d6	hepatitis c
708	30	81.1	256	2	Q6U5H9_ORYSA	Q6u5h9	oryza sativ
709	30	81.1	257	2	Q7U5GB1_ORYSA	Q7u5gb1	oryza sativ
710	30	81.1	258	2	Q4TFG9_TETNG	Q4tf99	tetraodon n
711	30	81.1	259	2	Q84CV7_9BACT	Q84cv7	uncultured
712	30	81.1	260	2	Q96IL7_HUMAN	Q96il7	homo sapien
713	30	81.1	260	2	Q7PYB1_ANOGA	Q7pyb1	anopheles g
714	30	81.1	262	1	SPIB_HUMAN	Q01892	homo sapien
715	30	81.1	262	2	Q5NVX0_9ARCH	Q5nvx0	uncultured
716	30	81.1	262	2	Q5Z3C2_NOCFA	Q5z3c2	nocardia fa
717	30	81.1	262	2	Q9JM46_RAT	Q9j4m6	rattus norv
718	30	81.1	262	2	Q6UDZ4_HCMV	Q6udz4	human cytom
719	30	81.1	264	2	Q96C55_HUMAN	Q96c55	homo sapien
720	30	81.1	264	2	Q6NW31_HUMAN	Q6nw31	homo sapien
721	30	81.1	266	2	Q52MA4_XENTR	Q52ma4	xenopus tro
722	30	81.1	267	2	Q6ZMK6_HUMAN	Q6zmk6	homo sapien
723	30	81.1	268	2	Q8T0C9_DROME	Q8t0c9	drosophila
724	30	81.1	268	2	Q6N9Z3_RHOPA	Q6n9z3	rhodopseudo
725	30	81.1	269	2	Q4W9L6_ASPFU	Q4w9l6	aspergillus
726	30	81.1	269	2	Q54SR6_DICDI	Q54sr6	dictyosteli
727	30	81.1	269	2	Q6H7I4_ORYSA	Q6h7i4	oryza sativ
728	30	81.1	271	2	Q9I5J4_PSEAB	Q9i5j4	pseudomonas
729	30	81.1	272	2	Q8RVW1_MUSAC	Q8rvw1	musa acumin
730	30	81.1	272	2	Q88RF6_PSEPK	Q88rf6	pseudomonas
731	30	81.1	272	2	Q99D13_BHV4	Q99d13	bovine herp
732	30	81.1	275	2	Q62364_CABEL	Q62364	caenorhabdi
733	30	81.1	278	1	MA32_MOUSE	Q35658	mus musculus
734	30	81.1	278	2	Q8H769_WHEAT	Q8h769	tritricum ae
735	30	81.1	279	2	Q4QC22_LEIMA	Q4qcn2	leishmania
736	30	81.1	279	2	Q9RWR7_DEIRA	Q9rrw7	deinococcus
737	30	81.1	279	2	Q8R5L1_MOUSE	Q8r5l1	mus musculus
738	30	81.1	279	2	Q6IRS5_RAT	Q6irs5	rattus norv
739	30	81.1	281	2	Q75AGO_ASHGO	Q75ag0	ashbya gos
740	30	81.1	281	2	Q5DFB3_SCHJA	Q5dfb3	schistosoma
741	30	81.1	281	2	Q9VLX3_DROME	Q9v1x3	drosophila
742	30	81.1	281	2	Q45765_CABEL	Q45765	caenorhabdi
743	30	81.1	283	2	Q52W85_DINO	Q52w85	akashwo sa
744	30	81.1	294	2	Q6K337_ORYSA	Q6k337	oryza sativ
745	30	81.1	294	2	Q5FPH3_GLUOX	Q5fph3	gluconobact
746	30	81.1	294	2	Q73ZQ4_MYCPA	Q73zq4	mycobacteri
747	30	81.1	295	2	Q9FVR6_ARATH	Q9fvr6	arabidopsis
748	30	81.1	296	2	Q9M4F1_BRANA	Q9m4f1	brassica na
749	30	81.1	299	2	Q9M4F1_BRANA	Q4zlm2	pseudomonas
750	30	81.1	300	2	Q4ZLM2_PSESY	Q9csv8	mus musculus
751	30	81.1	301	2	Q9CSV8_MOUSE	Q949e0	oryza sativ
752	30	81.1	302	2	Q949E0_ORYSA	Q4l9g1	burkholderi
753	30	81.1	302	2	Q4LGG1_9BURK	Q8e9v8	shewanella
754	30	81.1	303	2	Q8E9V8_SHEON	Q9sv84	arabidopsis
755	30	81.1	304	1	NIP51_ARATH	Q58cx8	bos taurus
756	30	81.1	311	2	Q58CX8_BOVIN	Q9lhf0	arabidopsis
757	30	81.1	312	2	Q9LHF0_ARATH	Q82c21	streptomyce
758	30	81.1	314	2	Q82C21_STRAW	Q6r960	hyphomicrob
759	30	81.1	322	2	Q6R960_9RHIZ	Q7rzv6	neurospora
760	30	81.1	324	2	Q7RZV6_NEUCR	Q8ghy1	pseudomonas
761	30	81.1	329	2	Q8GHY1_PSERE		

762	30	81.1	329	2	Q76BK6_PAROL	Q76bk6	paralichthy
763	30	81.1	333	2	Q8KQV3_VIBCH	Q8kqv3	vibrio chol
764	30	81.1	335	2	Q9TM68_ALECA	Q9tm68	alexandrium
765	30	81.1	338	2	Q6ZGQ0_ORYSA	Q6z9q0	oryza sativ
766	30	81.1	343	2	Q9TM72_PROMC	Q9tm72	prorocentru
767	30	81.1	345	2	Q55DQ5_DICDI	Q55dq5	dictyosteli
768	30	81.1	346	2	Q82MX3_STRAW	Q82mx3	streptomyce
769	30	81.1	346	2	Q7NET1_GLOVI	Q7net1	gloeobacter
770	30	81.1	347	2	Q9TM69_ALETA	Q9tm69	alexandrium
771	30	81.1	347	2	Q8B3W0_9GAMA	Q8b3w0	porcine lym
772	30	81.1	348	2	Q5DFV0_SCHJA	Q5dfv0	schistosoma
773	30	81.1	350	2	Q5BBY4_EMENI	Q5bby4	aspergillus
774	30	81.1	350	2	Q7XHXJ4_WHEAT	Q7xhj4	tritricum ae
775	30	81.1	352	2	Q9VWY7_DROME	Q9vwy7	drosophila
776	30	81.1	354	2	Q5JJC2_PYRKO	Q5jjc2	pyrococcus
777	30	81.1	357	2	Q8RXX2_PRORE	Q8rx2	providencia
778	30	81.1	363	2	Q8YZ86_ANASP	Q8yz86	anabaena sp
779	30	81.1	365	2	Q9SZF4_ARATH	Q9szf4	arabidopsis
780	30	81.1	367	1	CD5R2_HUMAN	Q13319	homo sapien
781	30	81.1	367	2	Q4ZFW6_HUMAN	Q4zfw6	homo sapien
782	30	81.1	367	2	Q5XKD4_HUMAN	Q5xkd4	homo sapien
783	30	81.1	367	2	Q8LFA7_ARATH	Q8lfa7	arabidopsis
784	30	81.1	368	2	Q8GWS2_ARATH	Q8gw52	arabidopsis
785	30	81.1	368	2	Q6EJ98_ARATH	Q6ej98	arabidopsis
786	30	81.1	368	2	Q5K481_SILDC	Q5k481	silene dicl
787	30	81.1	368	2	Q744W8_MYCPA	Q744w8	mycobacteri
788	30	81.1	370	2	Q5K480_SILCU	Q5k480	silene cucu
789	30	81.1	370	2	Q5K482_SILDC	Q5k482	silene dicl
790	30	81.1	370	2	Q8YL37_ANASP	Q8yl37	anabaena sp
791	30	81.1	371	2	Q8LGL6_SILLA	Q8lg16	silene lati
792	30	81.1	372	2	Q5K484_SILDI	Q5k484	silene dioi
793	30	81.1	372	2	Q8LGL1_SILDI	Q8lg11	silene dioi
794	30	81.1	372	2	Q8LGL2_9CARY	Q8lg12	lychnis flo
795	30	81.1	372	2	Q8LGL7_SILLA	Q8lg17	silene lati
796	30	81.1	373	2	Q8LGL4_SILCU	Q8lg14	silene cucu
797	30	81.1	373	2	Q8RXXB3_SILCO	Q8rxb3	silene conf
798	30	81.1	373	2	Q4UPL5_XANCP	Q4upl5	xanthomonas
799	30	81.1	373	2	Q8P434_XANCP	Q8p434	xanthomonas
800	30	81.1	376	2	Q9U5J1_9TRYP	Q9u5j1	trypanosoma
801	30	81.1	376	2	Q8VWN5_SILLA	Q8vwn5	silene lati
802	30	81.1	376	2	Q72E67_DESVH	Q72e67	desulfovibr
803	30	81.1	380	2	Q4XYH3_PLACH	Q4xyh3	plasmodium
804	30	81.1	384	2	Q67W96_ORYSA	Q67w96	oryza sativ
805	30	81.1	386	2	Q7WYR3_RHILV	Q7wyr3	rhizobium 1
806	30	81.1	387	2	Q4P8Y3_USTMA	Q4p8y3	ustilago ma
807	30	81.1	391	2	Q97553_CERAE	Q97553	cercopithec
808	30	81.1	392	2	Q5E9Q1_BOVIN	Q5e9q1	bos taurus
809	30	81.1	392	2	Q6GR57_XENLA	Q6gr57	xenopus lae
810	30	81.1	393	2	Q7UXV3_RHOBA	Q7uxv3	rhodopirell
811	30	81.1	395	2	Q4H9L0_9DEIO	Q4h9l0	deinococcus
812	30	81.1	395	2	Q8YP52_ANASP	Q8yp52	anabaena sp
813	30	81.1	396	2	Q4WJ66_ASPFU	Q4wj66	aspergillus
814	30	81.1	397	2	Q8R0W1_MOUSE	Q8r0w1	mus musculus
815	30	81.1	398	2	Q6K6D7_ORYSA	Q6k6d7	oryza sativ
816	30	81.1	402	2	Q85717_STRCU	Q85717	streptomyce
817	30	81.1	402	2	Q829Q6_STRAW	Q829q6	streptomyce
818	30	81.1	403	2	Q4RV84_TETNG	Q4rv84	tetraodon n
819	30	81.1	405	2	Q8LDP1_ARATH	Q8ldf1	arabidopsis
820	30	81.1	405	2	Q94BZ6_ARATH	Q94bz6	arabidopsis
821	30	81.1	408	2	Q7X692_ORYSA	Q7x692	oryza sativ
822	30	81.1	411	2	Q5N397_SYNPF	Q5n397	synecococc
823	30	81.1	413	2	Q5YUP8_NOCFA	Q5yup8	nocardia fa
824	30	81.1	413	2	Q7SZU5_XENLA	Q7szu5	xenopus lae
825	30	81.1	413	2	Q8QGW6_XENLA	Q8qgw6	xenopus lae
826	30	81.1	413	2	Q90XS5_XENLA	Q90xs5	xenopus lae
827	30	81.1	419	2	Q9RD81_STRCO	Q9rd81	streptomyce
828	30	81.1	419	2	Q9RIT4_STRCO	Q9rit4	streptomyce
829	30	81.1	422	2	Q8VWN3_SILLA	Q8vwn3	silene lati
830	30	81.1	425	2	Q4V626_DROME	Q4v626	drosophila
831	30	81.1	425	2	Q8VWN2_SILLA	Q8vwn2	silene lati
832	30	81.1	425	2	Q9LLR5_JUNEF	Q9llr5	juncus effu
833	30	81.1	426	2	Q8VWN4_SILLA	Q8vwn4	silene lati
834	30	81.1	430	2	Q67N20_SYMTH	Q67n20	symbiobacte

835	30	81.1	431	2	Q8UI19	AGRT5	Q8uil9	agrobacteri	908	30	81.1	545	2	Q4P5G9	USTMA	Q4p5g9	ustilago ma
836	30	81.1	435	2	Q7D1D9	AGRT5	Q7dl19	agrobacteri	909	30	81.1	549	2	Q8GUJ8	ARATH	Q8guj8	arabidopsis
837	30	81.1	438	2	Q50YN2	ENTHI	Q50yn2	entamoeba h	910	30	81.1	553	2	Q6IG04	RAT	Q6ig04	rattus norv
838	30	81.1	440	2	Q8MRX2	DROME	Q8mrx2	drosophila	911	30	81.1	558	2	Q61J10	CAEBR	Q61j10	caenorhabdi
839	30	81.1	440	2	Q9V871	DROME	Q9v871	drosophila	912	30	81.1	559	2	Q4PDA0	USTMA	Q4pda0	ustilago ma
840	30	81.1	441	2	Q5DFB1	SCHJA	Q5dfb1	schistosoma	913	30	81.1	560	2	Q7NJI4	GLOVI	Q7nji4	gloeobacter
841	30	81.1	442	2	Q9VAL8	DROME	Q9val8	drosophila	914	30	81.1	561	2	Q4WAY7	ASPFU	Q4way7	aspergillus
842	30	81.1	442	2	Q88HG2	PSEPK	Q88hg2	pseudomonas	915	30	81.1	562	2	Q6GPL0	XENLA	Q6gpl0	xenopus lae
843	30	81.1	443	2	Q80T99	MOUSE	Q80t99	mus musculu	916	30	81.1	565	2	Q59ZZ2	CANAL	Q59zz2	candida alb
844	30	81.1	444	2	Q52596	BJRAJA	Q52596	bradyrhizob	917	30	81.1	566	2	Q6CDR6	YARLI	Q6cdr6	yarrowia li
845	30	81.1	447	2	Q5E9I4	BOVIN	Q5e9i4	bos taurus	918	30	81.1	566	2	Q8CAC0	MOUSE	Q8cac0	mus musculu
846	30	81.1	447	2	Q8B0D8	9HEPC	Q8b0d8	hepatitis c	919	30	81.1	567	2	Q5A044	CANAL	Q5a044	candida alb
847	30	81.1	447	2	Q8B0D9	9HEPC	Q8b0d9	hepatitis c	920	30	81.1	567	2	Q51BC8	ENTHI	Q51bc8	entamoeba h
848	30	81.1	449	2	Q6N0I1	RHOPA	Q6n0i1	rhodopsedo	921	30	81.1	567	2	Q9LPW5	ARATH	Q9lpw5	arabidopsis
849	30	81.1	453	2	Q5CHG1	CRYHO	Q5chg1	cryptospori	922	30	81.1	568	2	Q6ZN96	HUMAN	Q6zn96	homo sapien
850	30	81.1	455	2	Q69MP5	ORYSA	Q69mp5	oryza sativ	923	30	81.1	577	2	Q51IP0	MAGGR	Q51ip0	magnaporth
851	30	81.1	457	2	Q60WL5	CAEBR	Q60wl5	caenorhabdi	924	30	81.1	579	2	Q6Q138	BOVIN	Q6q138	bos taurus
852	30	81.1	458	2	Q9SY19	ARATH	Q9sy19	arabidopsis	925	30	81.1	579	2	Q8QZS3	MOUSE	Q8qzs3	mus musculu
853	30	81.1	461	2	Q4TB03	TETNG	Q4tb03	tetraodon n	926	30	81.1	579	2	Q76JQ2	RAT	Q76jq2	rattus norv
854	30	81.1	462	2	Q7Z735	HUMAN	Q7z735	homo sapien	927	30	81.1	579	2	Q6SWP7	HCMV	Q6swp7	human cytom
855	30	81.1	462	2	Q68A51	RALSO	Q68a51	ralstonia s	928	30	81.1	580	1	VIE2	HCMVA	P19893	human cytom
856	30	81.1	462	2	Q5LNV7	SILPO	Q5lnv7	silicibacte	929	30	81.1	580	2	Q51UJ2	MAGGR	Q51uj2	magnaporth
857	30	81.1	462	2	Q8XYF8	RALSO	Q8xyf8	ralstonia s	930	30	81.1	580	2	Q6SW29	HCMV	Q6sw29	human cytom
858	30	81.1	463	2	Q51BD2	ENTHI	Q51bd2	entamoeba h	931	30	81.1	580	2	Q6SWJ2	HCMV	Q6swj2	human cytom
859	30	81.1	463	2	Q6N6F9	RHOPA	Q6n6f9	rhodopsedo	932	30	81.1	580	2	Q6SWL9	HCMV	Q6swl9	human cytom
860	30	81.1	464	2	Q9KHK2	9ACTO	Q9khk2	streptomyce	933	30	81.1	580	2	Q6SWS5	HCMV	Q6sws5	human cytom
861	30	81.1	466	2	Q51AU5	ENTHI	Q51au5	entamoeba h	934	30	81.1	580	2	Q6SWY2	HCMV	Q6swy2	human cytom
862	30	81.1	469	2	Q5EIT5	LISSE	Q5eit5	listeria se	935	30	81.1	580	2	Q6SX11	HCMV	Q6sx11	human cytom
863	30	81.1	469	2	Q8VMW8	LISSE	Q8vmw8	listeria se	936	30	81.1	580	2	Q6SX40	HCMV	Q6sx40	human cytom
864	30	81.1	469	2	Q9A3Y2	CAUCR	Q9a3y2	caulobacter	937	30	81.1	580	2	Q6SWV3	HCMV	Q6swv3	human cytom
865	30	81.1	470	2	Q6VMG4	BDEBA	Q6vmg4	bdellovibri	938	30	81.1	582	2	Q98683	9BETA	Q98683	simian cyto
866	30	81.1	475	2	Q6PKN2	9STRA	Q6pkn2	nitzschia c	939	30	81.1	586	2	Q8NA33	HUMAN	Q8na33	homo sapien
867	30	81.1	476	2	Q9V7E4	DROME	Q9v7e4	drosophila	940	30	81.1	595	2	Q73S88	MYCPA	Q73s88	mycobacteri
868	30	81.1	476	2	Q9V7E4	DROME	Q9v7e4	drosophila	941	30	81.1	595	2	Q82E16	STRAW	Q82e16	streptomyce
869	30	81.1	476	2	Q9ARH8	LYCES	Q9arh8	lycopersico	942	30	81.1	598	2	Q8QRY7	9BETA	Q8qry7	pongine her
870	30	81.1	477	1	MAZ	HUMAN	P56270	homo sapien	943	30	81.1	600	2	Q6AZT2	XENLA	Q6azt2	xenopus lae
871	30	81.1	477	1	MAZ	MOUSE	P56671	mus musculu	944	30	81.1	603	2	Q74DH2	GEOSL	Q74dh2	geobacter s
872	30	81.1	477	2	Q8M939	9ASTE	Q8m939	phyllonoma	945	30	81.1	603	2	Q9PTI1	XENLA	Q9pti1	xenopus lae
873	30	81.1	478	2	Q4I7F7	GIBZE	Q4i7f7	gibberella	946	30	81.1	607	2	Q4Q7G9	LEIMA	Q4q7g9	leishmania
874	30	81.1	478	2	Q4WJN9	ASPFU	Q4wjn9	aspergillus	947	30	81.1	609	2	Q94BT6	ARATH	Q94bt6	arabidopsis
875	30	81.1	479	2	Q9XGI8	LYCES	Q9xgi8	lycopersico	948	30	81.1	609	2	Q9LDF6	ARATH	Q9ldf6	arabidopsis
876	30	81.1	480	2	Q9ARH7	CAPAN	Q9arh7	capsicum an	949	30	81.1	610	2	Q7QTH7	GIALA	Q7qth7	giardia lam
877	30	81.1	480	2	Q39443	CAPAN	Q39443	capsicum an	950	30	81.1	612	2	Q8RUM5	ORYSA	Q8rum5	oryza sativ
878	30	81.1	481	2	Q8CXN8	OCEIH	Q8cxn8	oceanobacil	951	30	81.1	612	2	Q941S9	ORYSA	Q941s9	oryza sativ
879	30	81.1	482	2	Q59L98	CANAL	Q59l98	candida alb	952	30	81.1	614	1	TUP11	SCHPO	Q09715	schizosacch
880	30	81.1	488	2	Q89WD2	BJRAJA	Q89wd2	bradyrhizob	953	30	81.1	615	2	Q75DI6	ASHGO	Q75di6	ashbya goss
881	30	81.1	488	2	Q6UAQ1	TETNG	Q6uaq1	tetraodon n	954	30	81.1	615	2	Q6UEI5	MESCR	Q6uei5	mesembryant
882	30	81.1	490	2	Q82B66	STRAW	Q82b66	streptomyce	955	30	81.1	621	2	Q4RM99	TETNG	Q4rm99	tetraodon n
883	30	81.1	491	2	Q5NXH1	AZOSE	Q5nxxh1	azoarcus sp	956	30	81.1	622	2	Q617J8	CAEBR	Q617j8	caenorhabdi
884	30	81.1	492	1	CP26A	CHICK	Q9pub4	gallus gall	957	30	81.1	623	2	Q6BUJ5	DEBHA	Q6buj5	debaryomyce
885	30	81.1	493	2	Q8NFN7	HUMAN	Q8nfn7	homo sapien	958	30	81.1	623	2	Q8WZA9	HUMAN	Q8wza9	homo sapien
886	30	81.1	496	2	Q93YF8	9SOLA	Q93yf8	nicotiana a	959	30	81.1	624	1	SIR	SYNP7	P30008	synechococc
887	30	81.1	500	2	Q4WC97	ASPFU	Q4wc97	aspergillus	960	30	81.1	626	2	Q6Q079	CAEBR	Q60q79	caenorhabdi
888	30	81.1	500	2	Q24348	SORBI	Q24348	sorghum bic	961	30	81.1	628	2	Q4IR92	GIBZE	Q4ir92	gibberella
889	30	81.1	502	1	AKA80	RAT	Q35817	rattus norv	962	30	81.1	630	2	Q5Z8K3	ORYSA	Q5z8k3	oryza sativ
890	30	81.1	502	2	Q4PJW3	BOVIN	Q4pjwt3	bos taurus	963	30	81.1	630	2	Q8PMC2	XANAC	Q8pmc2	xanthomonas
891	30	81.1	502	2	Q7UE37	RHOPA	Q7ue37	rhodopirell	964	30	81.1	633	2	Q5N202	SYNP6	Q5n202	synechococc
892	30	81.1	503	2	Q4WJU0	ASPFU	Q4wjju0	aspergillus	965	30	81.1	634	1	PHAR2	HUMAN	Q75167	homo sapien
893	30	81.1	506	2	Q4I3L5	GIBZE	Q4i3l5	gibberella	966	30	81.1	634	2	Q5TFA0	HUMAN	Q5tfa0	homo sapien
894	30	81.1	509	2	Q4IBZ6	GIBZE	Q4ibz6	gibberella	967	30	81.1	635	2	Q5LS83	SILPO	Q5ls83	silicibacte
895	30	81.1	509	2	Q60979	LEIMA	Q60979	leishmania	968	30	81.1	637	2	Q5I0G4	RAT	Q5i0g4	rattus norv
896	30	81.1	510	1	CX56	CHICK	P29415	gallus gall	969	30	81.1	641	2	Q17740	CAEEL	Q17740	caenorhabdi
897	30	81.1	517	2	Q7S9S1	NEUCR	Q7s9s1	neurospora	970	30	81.1	641	2	Q7ZZW2	PAROL	Q7zzw2	paralichthy
898	30	81.1	519	2	Q4IJF2	GIBZE	Q4ijf2	gibberella	971	30	81.1	642	2	Q6BQJ8	DEBHA	Q6bqj8	debaryomyce
899	30	81.1	522	2	Q6PAZ2	XENLA	Q6paz2	xenopus lae	972	30	81.1	642	2	Q4YN59	PLABE	Q4yn59	plasmodium
900	30	81.1	524	2	Q5AWB5	EMENI	Q5awb5	aspergillus	973	30	81.1	643	2	Q7TYP4	MYCBO	Q7typ4	mycobacteri
901	30	81.1	527	2	Q75191	HUMAN	Q75191	homo sapien	974	30	81.1	643	2	P71750	MYCTU	P71750	mycobacteri
902	30	81.1	532	2	Q9IEE9	9HIV1	Q9iee9	human immun	975	30	81.1	644	2	Q8DLY7	SYNEL	Q8dly7	synechococc
903	30	81.1	534	2	Q8U3K6	PYRFU	Q8u3k6	pyrococcus	976	30	81.1	645	2	Q4WQ81	ASPFU	Q4wq81	aspergillus
904	30	81.1	536	2	Q7V997	PROMM	Q7v997	prochloroco	977	30	81.1	645	2	Q6PJ32	HUMAN	Q6pj32	homo sapien
905	30	81.1	541	1	GATB	YEAST	P33893	saccharomyc	978	30	81.1	647	2	Q5Y253	CRYGA	Q5y253	cryptococcu
906	30	81.1	542	2	Q09594	CABEL	Q09594	caenorhabdi	979	30	81.1	650	2	Q8YX67	ANASP	Q8yx67	anabaena sp
907	30	81.1	544	2	Q7MT55	WOLSU	Q7m7t5	wolinella s	980	30	81.1	650	2	Q58EJ8	BRARE	Q58ej8	brachydanio

```

981 30 81.1 652 2 Q7R4B0 GIALA
982 30 81.1 652 2 Q5N4E9_SYN6
983 30 81.1 653 2 Q5QG55_CRAGI
984 30 81.1 653 2 Q6GNY5_XENLA
985 30 81.1 659 2 Q5AXZ0_EMENI
986 30 81.1 660 2 Q4PGD1_USTMA
987 30 81.1 664 2 Q4GZ71_9TRYP
988 30 81.1 666 2 Q5FSZ2_GLUOX
989 30 81.1 670 2 Q593S5_CYPCA
990 30 81.1 672 2 Q613K3_CAEBR
991 30 81.1 695 2 Q8ZUD2_PYRAE
992 30 81.1 695 2 Q7SA97_NEUCR
993 30 81.1 697 1 Y4YR_RHISN
994 30 81.1 697 2 Q8RQEO_RHIFR
995 30 81.1 697 2 Q7ZYZ9_BRARE
996 30 81.1 703 1 CASC3_HUMAN
997 30 81.1 705 2 Q55UT6_CRYNE
998 30 81.1 705 2 Q5KHK2_CRYNE
999 30 81.1 708 2 Q8N1G5_HUMAN
1000 30 81.1 708 2 Q57ZF9_9TRYP

Q7r4b0 giardia lam
Q5n4e9 synechococc
Q5qgy5 crassostrea
Q6gny5 xenopus lae
Q5axz0 aspergillus
Q4pgd1 ustilago ma
Q4gz71 trypanosoma
Q5fsz2 gluconobact
Q593s5 cyprinus ca
Q613k3 caenorhabdi
Q8zud2 pyrobaculum
Q7sa97 neurospora
P55726 rhizobium s
Q8rqbo rhizobium f
Q7zyz9 brachydanio
O15234 homo sapien
Q55ut6 cryptococcu
Q5khk2 cryptococcu
Q8n1g5 homo sapien
Q57zf9 trypanosoma
```

ALIGNMENTS

```

RESULT 1
Q5MZB3_SYN6
ID Q5MZB3_SYN6 PRELIMINARY; PRT; 221 AA.
AC Q5MZB3;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=sync2417_d;
OS Synechococcus sp. (strain_PCC 6301) (Anacystis nidulans).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=269084;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PCC6301;
RA Sugita M.;
RT "Complete genome structure of the unicellular cyanobacterium Anacystis
RT nidulans 6301 (Synechococcus sp. PCC6301).";
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP008231; BAD80607.1; -; Genomic_DNA.
DR InterPro; IPR005829; Sug_transporter.
DR PROSITE; PS00217; SUGAR_TRANSPORT 2; UNKNOWN_1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 221 AA; 24058 MW; FAA89192DCEE750 CRC64;
```

```

Query Match 100.0%; Score 37; DB 2; Length 221;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 ILAPPVP 7
    |||||
DB 122 ILAPPVP 128
```

```

RESULT 2
Q9LGV4_ORYSA
ID Q9LGV4_ORYSA PRELIMINARY; PRT; 370 AA.
AC Q9LGV4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein P0706B05.32.
GN Name=P0706B05.32;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
```

```

RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Nilmura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
RA Ikeno M., Itoh S., Itoh T., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,
RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohta I., Ono N., Saij S., Sakai K., Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA Yano M., Jiang J., Gojobori T.;
RT "The genome sequence and structure of rice chromosome 1.";
RL Nature 420:312-316(2002).
DR EMBL; AP002482; BAA96629.1; -; Genomic_DNA.
DR Gramene; Q9LGV4; -.
KW Hypothetical protein.
SQ SEQUENCE 370 AA; 38964 MW; D9797D5B41FF04C9 CRC64;
```

```

Query Match 100.0%; Score 37; DB 2; Length 370;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 ILAPPVP 7
    |||||
DB 105 ILAPPVP 111
```

```

RESULT 3
Q8SXB2_DROME
ID Q8SXB2_DROME PRELIMINARY; PRT; 1064 AA.
AC Q8SXB2;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE GH15539p.
GN Name=Ptp4B; ORFNames=CG6899;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Berkelley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
DR EMBL; AY094714; AAM11067.1; -; mRNA.
DR HSSP; P10586; 1LAR.
DR FlyBase; FBgn0004368; CG6899.
DR FlyBase; FBgn0004368; Ptp4E.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004725; F:protein tyrosine phosphatase activity; IEA.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR InterPro; IPR003962; FNIIL_subd.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR000387; TYR_phosphatase.
DR InterPro; IPR000242; Tyr_PP.
DR Pfam; PF00041; fn3; 6.
DR Pfam; PF00102; Y_phosphatase; 1.
DR PRINTS; PR00014; FNTYPEPIII.
DR PRINTS; PR00700; PRTYPHPHTASE.
```

DR SMART; SM00060; FN3; 6.
DR SMART; SM00194; PTPC; 1.
DR PROSITE; PS50853; FN3; 6.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW Hydrolase; Repeat; Transmembrane.
SQ SEQUENCE 1064 AA; 120524 MW; 43F5DCBB8DEBAEC8 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 1064;
Best Local Similarity 100.0%; Pred. No. 7.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
|||
Db 553 ILAPPVP 559

RESULT 4
Q8IRS0 DROME PRELIMINARY; PRT; 1607 AA.
AC Q8IRS0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CG6899-PB, isoform B.
GN Name=Ptp4E; ORFNames=CG6899;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrial J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).

RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
a genomics perspective.";
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Milburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RG Berkeley Drosophila Genome Project;
RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
RA Yu C., Rubin G.;
RT "Drosophila melanogaster release 4 sequence.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE.
RG FlyBase;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
DR EMBL; AE003432; AAN09133.1; -; Genomic_DNA.
DR HSSP; P10586; 1LAR.
DR Ensembl; CG6899; Drosophila melanogaster.
DR FlyBase; FBgn0004368; CG6899.
DR FlyBase; FBgn0004368; Ptp4E.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004725; F:protein tyrosine phosphatase activity; IEA.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR InterPro; IPR003962; FNIII subd.
DR InterPro; IPR003961; FNIII.
DR InterPro; IPR000387; TYR_phosphatase.
DR InterPro; IPR000242; Tyr_PP.
DR Pfam; PF00041; fn3; 10.
DR Pfam; PF0102; Y_phosphatase; 1.
DR PRINTS; PR00014; FNTYPEIII.
DR PRINTS; PR00700; PRTYPHPHTASE.
DR SMART; SM00060; FN3; 11.
DR SMART; SM00194; PTPC; 1.
DR PROSITE; PS50853; FN3; 11.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW Hydrolase; Repeat; Transmembrane.
SQ SEQUENCE 1607 AA; 182285 MW; 6CAF554E61A27F45 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 1607;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
| | | | |
Db 1101 ILAPPVP 1107

RESULT 5

Q24495 DROME
ID Q24495_DROME PRELIMINARY; PRT; 1767 AA.
AC Q24495;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Receptor protein tyrosine phosphatase.
GN Name=Ptp4E;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=94043220; PubMed=82226938;
RA Oon S.H., Hong A., Yang X., Chia W.;
RT "Alternative splicing in a novel tyrosine phosphatase gene (DPTP4E) of
RT Drosophila melanogaster generates two large receptor-like proteins
RT which differ in their carboxyl termini.";
RL J. Biol. Chem. 268:23964-23971(1993).
DR EMBL; L20894; AAA76834.1; -; mRNA.
DR PIR; A49502; A49502.
DR HSSP; P18052; LYFO.
DR Ensembl; CG6899; Drosophila melanogaster.
DR FlyBase; FBgn0004368; Ptp4E.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004725; F:protein tyrosine phosphatase activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007275; P:development; IEA.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR003962; FN_III subd.
DR InterPro; IPR000387; TYR_phosphatase.
DR InterPro; IPR000242; Tyr_PP.
DR Pfam; PF00041; fn3; 10.
DR Pfam; PF00102; Y_phosphatase; 1.
DR PRINTS; PR00014; FNTYPEIII.
DR PRINTS; PR00700; PRTYPHPHTASE.
DR SMART; SM00060; FN3; 11.
DR SMART; SM00194; PTPC; 1.
DR PROSITE; PS50853; FN3; 11.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW Receptor.
SQ SEQUENCE 1767 AA; 200121 MW; 70E06692B8D4B89C CRC64;

Query Match 100.0%; Score 37; DB 2; Length 1767;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
| | | | |
Db 1101 ILAPPVP 1107

RESULT 6

Q9W4F5 DROME
ID Q9W4F5_DROME PRELIMINARY; PRT; 1767 AA.
AC Q9W4F5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CG6899-PA, isoform A.
GN Name=Ptp4E; ORFNames=CG6899;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20196006; PubMed=107311132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Goirell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskaas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
RT melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskaas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomics perspective.";
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RG Berkeley Drosophila Genome Project;
RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
RA Yu C., Rubin G.;
RT "Drosophila melanogaster release 4 sequence.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE.
RG FlyBase;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
RN [7]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=94043220; PubMed=8226938;
RA Oon S.H., Hong A., Yang X., Chia W.;
RT "Alternative splicing in a novel tyrosine phosphatase gene (DPTP4E) of
RT Drosophila melanogaster generates two large receptor-like proteins
RT which differ in their carboxyl termini.";
RL J. Biol. Chem. 268:23964-23971(1993).
CC -!- INTERACTION:
CC . Q9V6T4:CG13340; NbExp=1; IntAct=EBI-124740, EBI-98288;
CC Q9VQR9:CG17593; NbExp=1; IntAct=EBI-124740, EBI-99951;
CC Q9VX79:CG33067; NbExp=1; IntAct=EBI-124740, EBI-99044;
CC . Q9VFD1:CG6912; NbExp=1; IntAct=EBI-124740, EBI-102311;
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
DR EMBL; AE003432; AAF45998.1; -; Genomic_DNA.
DR PIR; A49502; A49502.
DR PIR; B49502; B49502.
DR HSSP; P10586; 1LAR.
DR IntAct; Q9W4F5; -.
DR Ensembl; CG6899; Drosophila melanogaster.
DR FlyBase; FBgn0004368; CG6899.
DR FlyBase; FBgn0004368; Ptp4E.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004725; F:protein tyrosine phosphatase activity; IEA.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR003962; FNIII subd.
DR InterPro; IPR00387; TYR_phosphatase.
DR InterPro; IPR00242; Tyr_pp.
DR Pfam; PF00041; fn3; 10.
DR Pfam; PF00102; Y_phosphatase; 1.
DR PRINTS; PR00014; FNTYPEIII.
DR PRINTS; PR00700; PRTYPHPHTASE.
DR SMART; SM00060; FN3; 11.
DR SMART; SM00194; PTPC; 1.
DR PROSITE; PS50853; FN3; 11.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW Hydrolase; Repeat; Transmembrane.
SQ SEQUENCE 1767 AA; 200075 MW; 4FF5CB358DB37B6C CRC64;

Query Match 100.0%; Score 37; DB 2; Length 1767;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
:|||||
Db 1101 ILAPPVP 1107

RESULT 7
Q39063 ARATH
ID Q39063_ARATH PRELIMINARY; PRT; 161 AA.
AC Q39063;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Cab3 promoter-binding protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Ecotype RLD;
RX MEDLINE=95241631; PubMed=7724680; DOI=10.1104/pp.107.2.657;
RA Mitra A., Higgins D.W.;
RT "Nucleotide sequence of a cDNA clone from Arabidopsis thaliana
RT encoding a small A-T tract-binding protein.";
RL Plant Physiol. 107:657-658(1995).
DR EMBL; L33781; AAA87293.1; -; mRNA.
SQ SEQUENCE 161 AA; 18479 MW; 0C1D81B7C9226B73 CRC64;

Query Match 97.3%; Score 36; DB 2; Length 161;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
:|||||
Db 123 VLAPPVP 129

RESULT 8
Q6ZN89 HUMAN
ID Q6ZN89_HUMAN PRELIMINARY; PRT; 218 AA.
AC Q6ZN89;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein FLJ16328.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Synovial membrane tissue;
RA Suzuki O., Sasaki N., Aotsuka S., Shoji T., Ichihara T., Shiohata N.,
RA Matsumoto K., Hirano M., Sano S., Nomura R., Yoshikawa Y.,
RA Matsumura Y., Moriya S., Chiba E., Momiyama H., Onogawa S.,
RA Kaeriyama S., Satoh N., Matsunawa H., Takahashi E., Kataoka R.,
RA Kuga N., Kuroda A., Satoh I., Kamata K., Takami S., Terashima Y.,
RA Watanabe M., Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A.,
RA Ishii S., Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Yamashita H., Matsuo K., Nakamura Y., Sekine M.,
RA Kikuchi H., Kanda K., Wagatsuma M., Murakawa K., Kanehori K.,
RA Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y.,
RA Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK131326; BAD18486.1; -; mRNA.
SQ SEQUENCE 218 AA; 23633 MW; 12D112E51063B4A8 CRC64;

Query Match 97.3%; Score 36; DB 2; Length 218;
Best Local Similarity 85.7%; Pred. No. 2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
:|||||
Db 93 VLAPPVP 99

RESULT 9

Q78BK3_NEUCR PRELIMINARY; PRT; 461 AA.
AC Q78BK3;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Predicted protein.
GN Name=NCU08547.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Iankiev P., Pedersen D., Nelson M., Washburne M.,
RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA Kamal M., Kamysseilis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
RA Krystofova S., Rasmussen C., Metznerberg R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catcheside D., Li W., Pratt R.J., Osmani S.A.,
RA DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Flammann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.,
RA "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL Nature 0:0-0(2003).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABX01000166; EAA33793.1; -; Genomic DNA.
SQ SEQUENCE 461 AA; 51895 MW; F6965A30AAF31089 CRC64;

Query Match 97.3%; Score 36; DB 2; Length 461;
Best Local Similarity 85.7%; Pred. No. 4.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7

Db 424 VLAPPVP 430

RESULT 10

Q628U5_ORYSA PRELIMINARY; PRT; 75 AA.
AC Q628U5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein P0686H11.1 (Hypothetical protein
DE P0605H02.41).
GN Name=P0686H11.1; Synonyms=P0605H02.41;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 8, PAC
clone:P0605H02.";

RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP004762; BAD09996.1; -; Genomic DNA.
DR EMBL; AP004620; BAD09674.1; -; Genomic DNA.

DR Gramene; Q628U5; -.

KW Hypothetical protein.

SQ SEQUENCE 75 AA; 8928 MW; 0542C39EECD9DF68 CRC64;

Query Match 94.6%; Score 35; DB 2; Length 75;
Best Local Similarity 71.4%; Pred. No. 98;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7

Db 64 VLAPPVP 70

RESULT 11

Q5D5E5_9RICK PRELIMINARY; PRT; 186 AA.
AC Q5D5E5;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Phenylacrylic acid decarboxylase, 3-octaprenyl-4-hydroxybenzoate
DE carboxy-lyase (EC 4.1.1.-).
GN Name=ubiX; ORFNames=WwMoj0053;
OS Wolbachia endosymbiont of Drosophila mojavensis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Wolbachieae; Wolbachia.
OX NCBI_TaxID=307503;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Salzberg S.L., Dunning Hotopp J.C., Delcher A.L., Pop M., Smith D.R.,
RA Eisen M.B., Nelson W.C.;
RT "Serendipitous discovery of Wolbachia genomes in multiple Drosophila
RT species.";
RL Genome Biol. 6:R23-R23(2005).
DR EMBL; AY897517; AAX14444.1; -; Genomic DNA.
DR InterPro; IPR003382; Flavoprotein.
DR InterPro; IPR004507; PAD1.
DR Pfam; PF02441; Flavoprotein; 1.
DR TIGRFAMs; TIGR00421; ubiX_pad; 1.
KW Lyase.
SQ SEQUENCE 186 AA; 20551 MW; 95B08191CBD0141E CRC64;

Query Match 94.6%; Score 35; DB 2; Length 186;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7

Db 144 ILAPPVP 150

RESULT 12

Q4UL63_RICFE PRELIMINARY; PRT; 191 AA.
AC Q4UL63;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE 3-octaprenyl-4-hydroxybenzoate carboxy-lyase (EC 4.1.1.-).
GN Name=ubiX; OrderedLocustNames=RF_0859;
OS Rickettsia felis (Rickettsia azadi).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsieae; Rickettsia; spotted fever group.
OX NCBI_TaxID=42862;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=URRWXCal2;
RX PubMed=15984913; DOI=10.1371/journal.pbio.0030248;
RA Ogata H., Renesto P., Audic S., Robert C., Blanc G., Fournier P.-E.,
RA Parinello H., Claverie J.-M., Raoult D.;

RT "The genome sequence of Rickettsia felis identifies the first putative
RT conjugative plasmid in an obligate intracellular parasite.";
RL PLoS Biol. 3:E248-E248(2005).
DR EMBL; CP000053; AAY61710.1; -; Genomic_DNA.
DR InterPro; IPR003382; Flavoprotein.
DR InterPro; IPR004507; PAD1.
DR Pfam; PF02441; Flavoprotein; 1.
DR TIGRFAMs; TIGR00421; ubiX_pad; 1.
KW Complete proteome; Lyase.
SQ SEQUENCE 191 AA; 21189 MW; 6D1C054513DE8582 CRC64;

Query Match 94.6%; Score 35; DB 2; Length 191;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
|:|||||
Db 146 IIAPPVP 152

RESULT 13
Q73HK2_WOLPM
ID Q73HK2_WOLPM PRELIMINARY; PRT; 191 AA.
AC Q73HK2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Phenylacrylic acid decarboxylase, 3-octaprenyl-4-hydroxybenzoate
carboxy-lyase (EC 4.1.1.-).
GN Name=ubiX; OrderedLocusNames=WD0556;
OS Wolbachia pipientis wMel.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Wolbachieae; Wolbachia.
OX NCBI_TaxID=66077;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15024419; DOI=10.1371/journal.pbio.0020069;
RA Wu M., Sun L.V., Vamathevan J.J., Riegler M., DeBoy R.T.,
RA Brownlie J.C., McGraw E.A., Martin W., Esser C., Ahmadinejad N.,
RA Wiegand C., Madupu R., Beanan M.J., Brinkac L.M., Daugherty S.C.,
RA Durkin A.S., Kolonay J.F., Nelson W.C., Mohamoud Y., Lee P.,
RA Berry K.J., Young M.B., Utterback T.R., Weidman J.F., Nierman W.C.,
RA Paulsen I.T., Nelson K.E., Tettelin H., O'Neill S.L., Eisen J.A.;
RT "Phylogenomics of the reproductive parasite Wolbachia pipientis wMel:
a streamlined genome overrun by mobile genetic elements.";
RL PLoS Biol. 2:327-341(2004).
DR EMBL; AE017258; AAS14263.1; -; Genomic_DNA.
DR TIGR; WD0556; -.
DR GO; GO:0016831; F:carboxy-lyase activity; IEA.
DR InterPro; IPR003382; Flavoprotein.
DR InterPro; IPR004507; PAD1.
DR Pfam; PF02441; Flavoprotein; 1.
DR TIGRFAMs; TIGR00421; ubiX_pad; 1.
KW Complete proteome; Lyase.
SQ SEQUENCE 191 AA; 21146 MW; FE4BB4EA173556DB CRC64;

Query Match 94.6%; Score 35; DB 2; Length 191;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
|:|||||
Db 144 IIAPPVP 150

RESULT 14
Q5PAP8_ANAMM
ID Q5PAP8_ANAMM PRELIMINARY; PRT; 201 AA.
AC Q5PAP8;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE 3-octaprenyl-4-hydroxybenzoate carboxyl-lyase (EC 4.1.1.-).

GN Name=ubiX; OrderedLocusNames=AM645;
OS Anaplasma marginale (strain St. Maries).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Anaplasmataceae; Anaplasma.
OX NCBI_TaxID=234826;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15618402; DOI=10.1073/pnas.040656102;
RA Brayton K.A., Kappmeyer L.S., Herndon D.R., Dark M.J., Tibbals D.L.,
RA Palmer G.H., McGuire T.C., Knowles D.P. Jr.;
RT "Complete genome sequencing of Anaplasma marginale reveals that the
RT surface is skewed to two superfamilies of outer membrane proteins.";
RL Proc. Natl. Acad. Sci. U.S.A. 102:844-849(2005).
DR EMBL; CP000030; AAV86632.1; -; Genomic_DNA.
DR GO; GO:0016831; F:carboxy-lyase activity; IEA.
DR GO; GO:0016829; F:lyase activity; IEA.
DR InterPro; IPR003382; Flavoprotein.
DR InterPro; IPR004507; PAD1.
DR Pfam; PF02441; Flavoprotein; 1.
DR TIGRFAMs; TIGR00421; ubiX_pad; 1.
KW Complete proteome; Lyase.
SQ SEQUENCE 201 AA; 22035 MW; 3FAFFA922641699A CRC64;

Query Match 94.6%; Score 35; DB 2; Length 201;
Best Local Similarity 85.7%; Pred. No. 2.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
|:|||||
Db 159 IIAPPVP 165

RESULT 15
Q60W03_CAEBR
ID Q60W03_CAEBR PRELIMINARY; PRT; 231 AA.
AC Q60W03;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG19323 (Fragment).
GN Name=CBG19323;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG The C.briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAC01000098; CAE72212.1; -; Genomic_DNA.
DR GO; GO:0005199; F:structural constituent of cell wall; IEA.
DR InterPro; IPR003882; Pistil_extensin.
DR PRINTS; PR01218; PSTLEXTENSIN.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 231 AA; 21940 MW; DC7DD1E86F46C959 CRC64;

Query Match 94.6%; Score 35; DB 2; Length 231;
Best Local Similarity 85.7%; Pred. No. 3.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
|:|||||
Db 112 IIAPPVP 118

RESULT 16
Q7D5X3_MYCTU
ID Q7D5X3_MYCTU PRELIMINARY; PRT; 266 AA.
AC Q7D5X3;

DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Soj family protein.
GN OrderedLocusNames=MT3309;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1773;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
DR EMBL; AE000516; AAK47651.1; -; Genomic_DNA.
DR TIGR; MT3309; -.
DR GO; GO:0042242; F:cobyrinic acid a,c-diamide synthase activity; IEA.
DR GO; GO:0009236; P:cobalamin biosynthesis; IEA.
DR InterPro; IPR002586; CbiA_P_synth.
DR Pfam; PF01656; CbiA; 1.
SQ SEQUENCE 266 AA; 28576 MW; 49BACF4E0ECC2116 CRC64;

Query Match 94.6%; Score 35; DB 2; Length 266;
Best Local Similarity 71.4%; Pred. No. 3.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
:|||||:
Db 209 VLAPPIP 215

RESULT 17
O05853_MYCTU
ID O05853_MYCTU PRELIMINARY; PRT; 266 AA.
AC O05853;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE POSSIBLE SOJ/PARA-RELATED PROTEIN.
GN OrderedLocusNames=RV3213c;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1773;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/311159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,
RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
DR EMBL; BX842582; CAB08303.1; -; Genomic_DNA.
DR PIR; F70595; F70595.
DR TubercuList; RV3213c; -.

DR GO; GO:0042242; F:cobyrinic acid a,c-diamide synthase activity; IEA.
DR GO; GO:0009236; P:cobalamin biosynthesis; IEA.
DR InterPro; IPR002586; CbiA_P_synth.
DR Pfam; PF01656; CbiA; 1.
KW Complete proteome.
SQ SEQUENCE 266 AA; 28575 MW; 419AC96E00CC2116 CRC64;

Query Match 94.6%; Score 35; DB 2; Length 266;
Best Local Similarity 71.4%; Pred. No. 3.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
:|||||:
Db 209 VLAPPIP 215

RESULT 18
Q73UQ2_MYCPA
ID Q73UQ2_MYCPA PRELIMINARY; PRT; 266 AA.
AC Q73UQ2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=MAP3314c;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017239; AAS05864.1; -; Genomic DNA.
DR GO; GO:0042242; F:cobyrinic acid a,c-diamide synthase activity; IEA.
DR GO; GO:0009236; P:cobalamin biosynthesis; IEA.
DR InterPro; IPR002586; CbiA_P_synth.
DR Pfam; PF01656; CbiA; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 266 AA; 28104 MW; 429E782A29C7781A CRC64;

Query Match 94.6%; Score 35; DB 2; Length 266;
Best Local Similarity 71.4%; Pred. No. 3.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
:|||||:
Db 209 VLAPPIP 215

RESULT 19
Q7TWZ9_MYCBO
ID Q7TWZ9_MYCBO PRELIMINARY; PRT; 266 AA.
AC Q7TWZ9;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE POSSIBLE SOJ/PARA-RELATED PROTEIN.
GN OrderedLocusNames=Mb3239c;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1765;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsempe C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.;

RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis."
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).
DR EMBL; BX248345; CAD95331.1; -; Genomic DNA.
DR GO; GO:0042242; P:cobyrinic acid a,c-diamide synthase activity; IEA.
DR GO; GO:0009236; P:cobalamin biosynthesis; IEA.
DR InterPro; IPR002586; CbiA_P_synth.
DR Pfam; PF01656; CbiA; 1.
KW Complete proteome.
SQ SEQUENCE 266 AA; 28576 MW; 49BACF4E0ECC2116 CRC64;

Query Match 94.6%; Score 35; DB 2; Length 266;
Best Local Similarity 71.4%; Pred. No. 3.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
Db 209 VLAPPIP 215

RESULT 20
Q5WMS2 ORYSA
ID Q5WMS2_ORYSA PRELIMINARY; PRT; 477 AA.
AC Q5WMS2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein OJ1333_C12.9.
GN Name=OJ1333_C12.9;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,
RA Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,
RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C., Leu H.-L.,
RA Li Y.-P., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
RA Wu H.-P., Shaw J.-F., *Yu Y., *Rambo T., *Currie J., *Collura K.,
RA *Soderlund C., *Wing R.;
RT "Oryza sativa BAC OJ1333_C12 genomic sequence."
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC104712; AAV32163.1; -; Genomic DNA.
DR InterPro; IPR003480; Transferase.
DR Pfam; PF02458; Transferase; 1.
KW Hypothetical protein.
SQ SEQUENCE 477 AA; 52057 MW; 44E7F43BD9637C96 CRC64;

Query Match 94.6%; Score 35; DB 2; Length 477;
Best Local Similarity 85.7%; Pred. No. 7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
Db 127 LLAPPVP 133

RESULT 21
Q7PYU0 ANOGA
ID Q7PYU0_ANOGA PRELIMINARY; PRT; 662 AA.
AC Q7PYU0;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000017547 (Fragment).
GN ORFNames=ENSANGG0000015058;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.

OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008987; EAA01467.1; -; Genomic_DNA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003684; F:damaged DNA binding; IEA.
DR GO; GO:0006289; P:nucleotide-excision repair; IEA.
DR InterPro; IPR002345; Lipocalin.
DR InterPro; IPR004583; Rad4.
DR PANTHER; PTHR12135; Rad4; 1.
DR Pfam; PF03835; Rad4; 1.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 662 AA; 76575 MW; 0B18D6974A185D87 CRC64;

Query Match 94.6%; Score 35; DB 2; Length 662;
Best Local Similarity 85.7%; Pred. No. 9.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
Db 141 ILAPPVP 147

RESULT 22
Q55ZY4 CRYNE
ID Q55ZY4_CRYNE PRELIMINARY; PRT; 842 AA.
AC Q55ZY4;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=CNBA3560;
OS Cryptococcus neoformans var. neoformans B-3501A.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX NCBI_TaxID=283643;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B-3501A;
RA Fung E., Hyman R.W., Rowley D., Bruno D., Miranda M., Fukushima M.,
RA Wickes B.L., Fu J., Davis R.W.;
RT "Cryptococcus neoformans serotype D sequencing."
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AA01000002; EAL23240.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 842 AA; 94227 MW; DA8051D58EC1ABE2 CRC64;

Query Match 94.6%; Score 35; DB 2; Length 842;
Best Local Similarity 71.4%; Pred. No. 1.3e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
Db 200 VLAPPIP 206

RESULT 23
Q5KP92 CRYNE
ID Q5KP92_CRYNE PRELIMINARY; PRT; 842 AA.
AC Q5KP92;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)

DE 1-phosphatidylinositol-4-phosphate 5-kinase, putative.
GN ORFNames=CNA03760;
OS Cryptococcus neoformans var. neoformans JEC21.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX NCBI_TaxID=214684;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=JEC21;
RA Loftus B.J., Fung E., Roncaglia P., Rowley D., Amedeo P., Bruno D.,
RA Vamathevan J., Miranda M., Anderson I.J., Fraser J.A., Allen J.E.,
RA Bosdet I.E., Brent M.R., Chiu R., Doering T.L., Donlin M.J.,
RA D'Souza C.A., Fox D.S., Grinberg V., Fu J., Fukushima M., Haas B.J.,
RA Huang J.C., Janbon G., Jones S.J.M., Krzywinski M.I., Kwon-Chung J.K.,
RA Lengeler K.B., Maiti R., Marra M.A., Marra R.E., Mathewson C.A.,
RA Mitchell T.G., Pertea M., Riggs F.R., Salzberg S.L., Shvartsbeyn A.,
RA Schein J.E., Shin H., Specht C.A., Suh B., Tenney A., Utterback T.,
RA Wickes B.L., Wye N.H., Kronstad J., Lodge J.K., Heitman J.,
RA Davis R.W., Fraser C.M., Hyman R.W.;
RT "The genome and transcriptome of Cryptococcus neoformans, a
RT basidiomycete fungal pathogen of humans.";
RL Science 0:0-0(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=JEC21;
RA Loftus B., Amedeo P., Roncaglia P., Vamathevan J., Utterback T.,
RA Van Aken S., Fraser C.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=JEC21;
RX PubMed=15653466; DOI=10.1126/science.1103773;
RA Loftus B.J., Fung E., Roncaglia P., Rowley D., Amedeo P., Bruno D.,
RA Vamathevan J., Miranda M., Anderson I.J., Fraser J.A., Allen J.E.,
RA Bosdet I.E., Brent M.R., Chiu R., Doering T.L., Donlin M.J.,
RA D'Souza C.A., Fox D.S., Grinberg V., Fu J., Fukushima M., Haas B.J.,
RA Huang J.C., Janbon G., Jones S.J.M., Koo H.L., Krzywinski M.I.,
RA Kwon-Chung K.J., Lengeler K.B., Maiti R., Marra M.A., Marra R.E.,
RA Mathewson C.A., Mitchell T.G., Pertea M., Riggs F.R., Salzberg S.L.,
RA Schein J.E., Shvartsbeyn A., Shin H., Shumway M., Specht C.A.,
RA Suh B.B., Tenney A., Utterback T.R., Wickes B.L., Wortman J.R.,
RA Wye N.H., Kronstad J.W., Lodge J.K., Heitman J., Davis R.W.,
RA Fraser C.M., Hyman R.W.;
RT "The genome of the basidiomycetous yeast and human pathogen
RT Cryptococcus neoformans.";
RL Science 307:1321-1324(2005).
DR EMBL; AE017341; AAW40890.1; -; Genomic DNA.
DR GO; GO:0016308; F:1-phosphatidylinositol-4-phosphate 5-kinase. . .; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR InterPro; IPR002498; PIP5K.
DR Pfam; PF01504; PIP5K; 1.
DR SMART; SM00330; PIPKc; 1.
KW Complete proteome; Kinase.
SQ SEQUENCE 842 AA; 94228 MW; DA8051D58EC1ABE2 CRC64;

Query Match 94.6%; Score 35; DB 2; Length 842;
Best Local Similarity 71.4%; Pred. No. 1.3e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
:||||:|
Db 200 VLAPPIP 206

RESULT 24
Q6Y8G9 LEIME
ID Q6Y8G9_LEIME PRELIMINARY; PRT; 990 AA.
AC Q6Y8G9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Beta adaptin.
OS Leishmania mexicana mexicana.

OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=44270;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MNYC/BZ/62/M379;
RA Gokool S.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY170466; AA017687.1; -; Genomic DNA.
DR InterPro; IPR002553; Adaptin_N.
DR InterPro; IPR002088; PPTA.
DR InterPro; IPR008162; Pyrophosphatase.
DR Pfam; PF01602; Adaptin_N; 1.
DR PROSITE; PS00387; PPASE; UNKNOWN_1.
DR PROSITE; PS00904; PPTA; UNKNOWN_1.
SQ SEQUENCE 990 AA; 108292 MW; 346B7BDB80F9A013 CRC64;

Query Match 94.6%; Score 35; DB 2; Length 990;
Best Local Similarity 71.4%; Pred. No. 1.5e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
:||||:|
Db 859 VLAPPIP 865

RESULT 25
Q4QGY0 LEIMA
ID Q4QGY0_LEIMA PRELIMINARY; PRT; 990 AA.
AC Q4QGY0;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Adaptin-related protein-like protein.
GN ORFNames=LmjF11.0990;
OS Leishmania major.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Friedlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neil S.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; CT005250; CAJ02880.1; -; Genomic DNA.
SQ SEQUENCE 990 AA; 108501 MW; 8D752523AEA39F7F CRC64;

Query Match 94.6%; Score 35; DB 2; Length 990;
Best Local Similarity 71.4%; Pred. No. 1.5e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
:||||:|
Db 860 VLAPPIP 866

RESULT 26
PAAD_RICPR
ID PAAD_RICPR STANDARD; PRT; 189 AA.
AC Q9ZD09;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Probable aromatic acid decarboxylase (EC 4.1.1.-).
GN OrderedLocusNames=RP541;
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia; typhus group.
OX NCBI_TaxID=782;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Madrid E;
RX MEDLINE=99039499; PubMed=9823893; DOI=10.1038/24094;

RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA Sicheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria";
RL Nature 396:133-140(1998).
CC -1- SIMILARITY: Belongs to the polyprenyl p-hydroxybenzoate /
CC phenylacrylic acid decarboxylases family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AJ235272; CAA14990.1; -; Genomic_DNA.
DR PIR; D71658; D71658.
DR InterPro; IPR003382; Flavoprotein.
DR InterPro; IPR004507; PAD1.
DR Pfam; PF02441; Flavoprotein; 1.
DR TIGRFAMs; TIGR00421; ubiX_pad; 1.
KW Complete proteome; Decarboxylase; Hypothetical protein; Lyase.
SQ SEQUENCE 189 AA; 20973 MW; DD2DBF374C30B8E6 CRC64;

Query Match 91.9%; Score 34; DB 1; Length 189;
Best Local Similarity 71.4%; Pred. No. 3.9e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
Db 146 IIAPPVP 152
|:|||||
"
RESULT 27
Q68WJ3 RICTY
ID Q68WJ3_RICTY PRELIMINARY; PRT; 190 AA.
AC Q68WJ3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE 3-octaprenyl-4-hydroxybenzoate carboxy-lyase (EC 4.1.1.-).
GN Name=ubiX; OrderedLocusNames=RT0530;
OS Rickettsia typhi.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsieae; Rickettsia; typhus group.
OX NCBI_TaxID=785;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Wilmington;
RX PubMed=15317790; DOI=10.1128/JB.186.17.5842-5855.2004;
RA McLeod M.P., Qin X., Karpathy S.E., Gioia J., Highlander S.K.,
RA Fox G.E., McNeill T.Z., Jiang H., Muzny D., Jacob L.S., Hawes A.C.,
RA Sodergren E., Gill R., Hume J., Morgan M., Fan G., Amin A.G.,
RA Gibbs R.A., Hong C., Yu X.-J., Walker D.H., Weinstock G.M.;
RT "Complete genome sequence of Rickettsia typhi and comparison with
RT sequences of other Rickettsiae";
RL J. Bacteriol. 186:5842-5855(2004).
DR EMBL; AE017197; AAU03999.1; -; Genomic_DNA.
DR GO; GO:0016829; F:lyase activity; IEA.
DR InterPro; IPR003382; Flavoprotein.
DR InterPro; IPR004507; PAD1.
DR Pfam; PF02441; Flavoprotein; 1.
DR TIGRFAMs; TIGR00421; ubiX_pad; 1.
KW Complete proteome; Lyase.
SQ SEQUENCE 190 AA; 20858 MW; B2E82B16DBDD2A67 CRC64;

Query Match 91.9%; Score 34; DB 2; Length 190;
Best Local Similarity 71.4%; Pred. No. 4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
Db 146 IIAPPVP 152
|:|||||

RESULT 28
Q62I56 BURMA
ID Q62I56_BURMA PRELIMINARY; PRT; 315 AA.
AC Q62I56;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Transcriptional regulator, LysR family.
GN OrderedLocusNames=BMA2032;
OS Burkholderia mallei (Pseudomonas mallei).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=13373;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 23344;
RX PubMed=15377793; DOI=10.1073/pnas.0403306101;
RA Nieman W.C., DeShazer D., Kim H.S., Tettelin H., Nelson K.E.,
RA Feldblyum T.V., Ulrich R.L., Ronning C.M., Brinkac L.M.,
RA Daugherty S.C., Davidsen T.D., DeBoy R.T., Dimitrov G., Dodson R.J.,
RA Durkin A.S., Gwinn M.L., Haft D.H., Khouri H.M., Kolonay J.F.,
RA Madupu R., Mohammed Y., Nelson W.C., Radune D., Romero C.M.,
RA Sarria S., Selengut J., Shamblin C., Sullivan S.A., White O., Yu Y.,
RA Zafar N., Zhou L., Fraser C.M.;
RT "Structural flexibility in the Burkholderia mallei genome.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14246-14251(2004).
CC -1- SIMILARITY: Contains 1 HTH LysR-type DNA-binding domain.
DR EMBL; CP000010; AAU49569.1; -; Genomic_DNA.
DR TIGR; BMA2032;
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR000847; HTH_LysR.
DR InterPro; IPR005119; LysR_subst_bd.
DR InterPro; IPR011991; Wing_hlx_DNA_bd.
DR Pfam; PF00126; HTH_1; 1.
DR Pfam; PF03466; LysR_substrate; 1.
DR PRINTS; PR00039; HTHLYSR.
DR PROSITE; PS50931; HTH_LYSR; 1.
KW Activator; Complete proteome; DNA-binding; Transcription;
KW Transcription regulation.
SQ SEQUENCE 315 AA; 34879 MW; 8BAD4282598CD429 CRC64;

Query Match 91.9%; Score 34; DB 2; Length 315;
Best Local Similarity 71.4%; Pred. No. 6.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
Db 268 VIAPPVP 274
|:|||||
"
RESULT 29
Q5RI92 BRARE
ID Q5RI92_BRARE PRELIMINARY; PRT; 341 AA.
AC Q5RI92;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Novel protein similar to prostaglandin receptor E2 (Subtype EP2)
DE (PTGER2).
GN Name=OTTDARPP00000003831; ORFNames=CH211-13N20.1-001;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Hunter G.;
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.

```
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
DR EMBL; BX284679; CAI11598.1; -; Genomic_DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0004960; F:thromboxane receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR00276; GPCR_Rhodp.
DR InterPro; IPR00370; IP_receptor.
DR InterPro; IPR008365; Prostanoidrecept.
DR InterPro; IPR01105; Thbox_receptor.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCRRHODOPSN.
DR PRINTS; PR01788; PROSTANOIDR.
DR PRINTS; PR00856; PRSTNOIDIPR.
DR PRINTS; PR00429; THROMBOXANER.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Receptor; Transducer; Transmembrane.
SQ SEQUENCE 341 AA; 38790 MW; CBP1C4D5DA69E1CD CRC64;

Query Match 91.9%; Score 34; DB 2; Length 341;
Best Local Similarity 85.7%; Pred. No. 7.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
Db 300 ILSPVP 306

RESULT 30
Q4SLJ8 TETNG
ID Q4SLJ8_TETNG PRELIMINARY; PRT; 467 AA.
AC Q4SLJ8;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 15 SCAF14556, whole genome shotgun sequence.
GN ORFNames=GSTENG0016208001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castellini V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Guzy J.,
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff JN., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Winkler P., Lander E.S., Weissbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAE01014556; CAF98484.1; -; Genomic DNA.
SQ SEQUENCE 467 AA; 49745 MW; 9517E9DB4F0E6D37 CRC64;
```

```
Query Match 91.9%; Score 34; DB 2; Length 467;
Best Local Similarity 71.4%; Pred. No. 1e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
Db 352 LLAPPVP 358

RESULT 31
Q4IM73 GIBZE
ID Q4IM73_GIBZE PRELIMINARY; PRT; 621 AA.
AC Q4IM73;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=FG01685.1;
OS Gibberella zeae PH-1.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
OX NCBI_TaxID=229533;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PH-1;
RA Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavkiy L.,
RA Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA Choepe Y., Collymore A., Cook A., Cooke P., Corum B., DeArelano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA Roman J., Schauer S., Schupback R., Seaman S., Severy P., Smirnov S.,
RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA Lander E.;
RT "Fusarium graminearum genome sequence.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACW0100090; EAA67830.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 621 AA; 70293 MW; 7CC8E3E25713A93B CRC64;

Query Match 91.9%; Score 34; DB 2; Length 621;
Best Local Similarity 71.4%; Pred. No. 1.4e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
Db 451 VIAPPVP 457

RESULT 32
Q8T3T1 LYTVA
ID Q8T3T1_LYTVA PRELIMINARY; PRT; 762 AA.
AC Q8T3T1;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Major yolk protein (Fragment).
```


OS	Lytechinus variegatus (Sea urchin).				
OC	Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;				
OC	Echinoidea; Euechinoidea; Echinacea; Temnopleuroidea; Toxopneustidae;				
OC	Lytechinus.				
OX	NCBI_TaxID=7654;				
RN	[1]				
RP	NUCLEOTIDE SEQUENCE.				
RX	MEDLINE=21965081; PubMed=11969251; DOI=10.1006/dbio.2002.0611;				
RA	Brooks J.M., Wessel G.M.;				
RT	"The major yolk protein in sea urchins is a transferrin-like, iron				
RT	binding protein.";				
RL	Dev. Biol. 245:1-12(2002).				
DR	EMBL; AY090113; AAM14718.1; -; mRNA.				
DR	HSSP; P56410; LOVB.				
DR	GO; GO:0005576; C:extracellular region; IEA.				
DR	GO; GO:0008199; F:ferric iron binding; IEA.				
DR	GO; GO:0006879; P:iron ion homeostasis; IEA.				
DR	GO; GO:0006826; P:iron ion transport; IEA.				
DR	GO; GO:0006810; P:transport; IEA.				
DR	InterPro; IPR001156; Peptidase_S60.				
DR	InterPro; IPR008197; WAP.				
DR	Pfam; PF00405; Transferrin; 1.				
DR	Pfam; PF00095; WAP; 1.				
DR	PRINTS; PR00422; TRANSFERRIN.				
DR	SMART; SM00094; TR FER; 1.				
DR	SMART; SM00217; WAP; 1.				
FT	NON TER 762 762				
SQ	SEQUENCE 762 AA; 86711 MW; BEE1B4EA11CA3373 CRC64;				
Query Match 91.9%; Score 34; DB 2; Length 762;					
Best Local Similarity 85.7%; Pred. No. 1.7e+03;					
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;					
QY	1 ILAPPVP 7				
Db	756 ILSPPVP 762				
RESULT 33					
Q8C585	MOUSE				
ID	Q8C585_MOUSE PRELIMINARY; PRT; 768 AA.				
AC	Q8C585;				
DT	01-MAR-2003 (TrEMBLrel. 23, Created)				
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)				
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)				
DE	Mus musculus adult male urinary bladder cDNA, RIKEN full-length				
DE	enriched library, clone:9530051E16 product:weakly similar to PERQ1				
DE	(Fragment).				
GN	Name=Tnrc15;				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;				
OC	Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	NUCLEOTIDE SEQUENCE.				
RC	STRAIN=C57BL/6J; TISSUE=Urinary bladder;				
RX	MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;				
RA	Carninci P., Hayashizaki Y.;				
RT	"High-efficiency full-length cDNA cloning.";				
RL	Meth. Enzymol. 303:19-44(1999).				
RN	[2]				
RP	NUCLEOTIDE SEQUENCE.				
RC	STRAIN=C57BL/6J; TISSUE=Urinary bladder;				
RX	MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;				
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,				
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,				
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,				
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,				
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,				
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,				
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,				
RA	Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,				

RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,				
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,				
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,				
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,				
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,				
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,				
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,				
RA	Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,				
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,				
RA	Hayashizaki Y.;				
RT	"Functional annotation of a full-length mouse cDNA collection.";				
RL	Nature 409:685-690(2001).				
RN	[3]				
RP	NUCLEOTIDE SEQUENCE.				
RC	STRAIN=C57BL/6J; TISSUE=Urinary bladder;				
RX	MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;				
RA	Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,				
RA	Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;				
RT	"Normalization and subtraction of cap-trapper-selected cDNAs to				
RT	prepare full-length cDNA libraries for rapid discovery of new genes.";				
RL	Genome Res. 10:1617-1630(2000).				
RN	[5]				
RP	NUCLEOTIDE SEQUENCE.				
RC	STRAIN=C57BL/6J; TISSUE=Urinary bladder;				
RX	MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;				
RA	Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,				
RA	Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,				
RA	Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,				
RA	Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,				
RA	Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,				
RA	Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,				
RA	Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;				
RT	"RIKEN integrated sequence analysis (RISA) system-384-format				
RT	sequencing pipeline with 384 multicapillary sequencer.";				
RL	Genome Res. 10:1757-1771(2000).				
RN	[6]				
RP	NUCLEOTIDE SEQUENCE.				
RC	STRAIN=C57BL/6J; TISSUE=Urinary bladder;				
RA	Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,				
RA	Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,				
RA	Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,				
RA	Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,				
RA	Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,				
RA	Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,				
RA	Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,				
RA	Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,				
RA	Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,				
RA	Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,				
RA	Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;				
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AK079246; BAC37586.1; -; mRNA.				
DR	Ensembl; ENSMUSG00000048000; Mus musculus.				
DR	MGI; MGI:2138584; Tnrc15.				
DR	InterPro; IPR003169; GYF.				
DR	Pfam; PF02213; GYF; 1.				
DR	SMART; SM00444; GYF; 1.				
DR	PROSITE; PS50829; GYF; 1.				
FT	NON TER 768 768				
SQ	SEQUENCE 768 AA; 87239 MW; 54FA470DF3653D57 CRC64;				

Query Match 91.9%; Score 34; DB 2; Length 768;					
Best Local Similarity 85.7%; Pred. No. 1.7e+03;					
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;					
QY	1 ILAPPVP 7				

Db 445 ILSPVP 451
|||||

```
RESULT 34
P73172_Syny3
ID P73172_Syny3 PRELIMINARY; PRT; 924 AA.
AC P73172;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CheA like protein.
GN OrderedLocusNames=s111296;
OC Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugiyura M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
DR EMBL; BA000022; BAA17198.1; -; Genomic_DNA.
DR PIR; S75284; S75284.
DR HSSP; Q56310; 1158.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0000156; F:two-component response regulator activity; IEA.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0006935; P:chemotaxis; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR002545; Chew.
DR InterPro; IPR004105; H-kinase dim.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR004358; His_kin_like_C.
DR InterPro; IPR008207; Hpt.
DR InterPro; IPR008208; Hpt_N.
DR InterPro; IPR001789; Response_reg.
DR Pfam; PF01584; Chew; 1.
DR Pfam; PF02895; H-kinase_dim; 1.
DR Pfam; PF02518; HATPase_C; 1.
DR Pfam; PF01627; Hpt; 1.
DR Pfam; PF00072; Response_reg; 1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR ProDom; PD003142; Hpt_N; 1.
DR SMART; SM000039; Response_reg; 1.
DR SMART; SM00260; Chew; 1.
DR SMART; SM00387; HATPase_C; 1.
DR SMART; SM00073; HPT; 1.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS50109; HIS_KIN; 1.
DR PROSITE; PS50894; HPT; 1.
DR PROSITE; PS50110; RESPONSE_REGULATORY; 1.
KW Complete proteome.
SQ SEQUENCE 924 AA; 101710 MW; 6A70A7931C814588 CRC64;
```

```
Query Match 91.9%; Score 34; DB 2; Length 924;
Best Local Similarity 71.4%; Pred. No. 2.1e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
:|||||
Db 279 VIAPPVP 285
```

```
RESULT 35
Q520P7_MAGGR PRELIMINARY; PRT; 959 AA.
AC Q520P7;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=MG05303.4;
OS Magnaporthe grisea 70-15.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.
OX NCBI_TaxID=242507;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
RA Bayul T., Blitshsteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collamore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Grandbois E., Gyaltsen K., Hafez N.,
RA Gnirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamvyselis M., Karlsson E.,
RA Kells C., Kieu A., Kisner P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Maucelli E.,
RA Mccarthy M., Mcdonough S., Mcghee T., Meldrim J., Meneus L.,
RA Mesirov J., Mihalev A., Mihova T., Mikkelson T., Mlenga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., Neues C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okaowo O., O'leary S., Omotosho B.,
RA O'neill K., Osman S., Parker S., Perrin D., Phunkhang P., Piquani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thoulutsang Y., Topham K.,
RA Towey S., Tsamla T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Magnaporthe grisea.";
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
[2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Dean R., Mitchell T., Brown D., Pan H., Thon M.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
[3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Zhu H., Blackmon B.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACU01000810; EAA52611.1; -; Genomic_DNA.
KW Hypothetical protein.
```

SQ SEQUENCE 959 AA; 106289 MW; ADCE86F580735896 CRC64;

Query Match 91.9%; Score 34; DB 2; Length 959;
Best Local Similarity 71.4%; Pred. No. 2.2e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
: : : : :
Db 433 VMAPPVP 439

RESULT 36
Q52FV3_MAGGR
ID Q52FV3_MAGGR PRELIMINARY; PRT; 1011 AA.
AC Q52FV3;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=MG01533.4;
OS Magnaporthe grisea 70-15.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.
OX NCBI_TaxID=242507;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
RA Bayul T., Blitshsteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collimore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA Gnirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamysseles M., Karlsson E.,
RA Kells C., Kieu A., Kisner P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., Mcghee T., Meldrim J., Meneus L.,
RA Mesirov J., Mihalev A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., Newes C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
RA O'neill K., Osman S., Parker S., Perrin D., Phunkhang P., Pigani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Testaye S., Theodore J., Thoulutsang Y., Topham K.,
RA Towey S., Tsamla T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whitaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Magnaporthe grisea.";
RL Submitted (OCT-2003) to the EMBL/GenBank/DBSJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Dean R., Mitchell T., Brown D., Pan H., Thon M.;
RA Submitted (OCT-2003) to the EMBL/GenBank/DBSJ databases.
RN [3]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Zhu H., Blackmon B.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBSJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBSJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACU01000067; EAA55882.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 1011 AA; 107576 MW; 81A6BBA934D950AB CRC64;

Query Match 91.9%; Score 34; DB 2; Length 1011;
Best Local Similarity 85.7%; Pred. No. 2.3e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
: : : : :
Db 139 ILAPPLP 145

RESULT 37
Q80TV1_MOUSE
ID Q80TV1_MOUSE PRELIMINARY; PRT; 1048 AA.
AC Q80TV1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE MKIAA0642 protein (Fragment).
GN Name=Trnc15; Synonyms=mKIAA0642;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=22579291; PubMed=12693553;
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,
RA Nakajima D., Nagase T., Ohara O., Koga H.;
RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
RT II. The complete nucleotide sequences of 400 mouse KIAA-homologous
RT cDNAs identified by screening of terminal sequences of cDNA clones
RT randomly sampled from size-fractionated libraries.";
RL DNA Res. 10:35-48(2003).
DR EMBL; AK122337; BAC65619.1; -; mRNA.
DR HSSP; Q05195; 1NLW.
DR Ensembl; ENSMUSG0000048000; Mus musculus.
DR MGI; MGI:2138584; Tnrc15.
DR InterPro; IPR003169; GYF.
DR Pfam; PF02213; GYF; 1.
DR SMART; SM00444; GYF; 1.
DR PROSITE; PSS0829; GYF; 1.
FT NON_TER 1
SQ SEQUENCE 1048 AA; 121839 MW; 113C516E81E98A4C CRC64;

Query Match 91.9%; Score 34; DB 2; Length 1048;
Best Local Similarity 85.7%; Pred. No. 2.4e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
: : : : :
Db 208 ILSPFPV 214

RESULT 38
Q6Y7W8_MOUSE
ID Q6Y7W8_MOUSE PRELIMINARY; PRT; 1291 AA.
AC Q6Y7W8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Grb10 interacting GYF protein 2.


```
GN Name=Tnrc15;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=22803289; PubMed=12771153; DOI=10.1074/jbc.M211572200;
RA Giovannone B., Lee E., Laviola L., Giorgino F., Cleveland K.A.,
RA Smith R.J.;
RT "Two novel proteins that are linked to insulin-like growth factor
RL (IGF-I) receptors by the Grl0 adapter and modulate IGF-I signaling.";
RL J. Biol. Chem. 278:31564-31573 (2003).
DR EMBL; AY176043; AAO46887.1; -; mRNA.
DR MGI; MGI:2138584; Tnrc15.
DR InterPro; IPR003169; GYF.
DR Pfam; PF02213; GYF; 1.
DR SMART; SM00444; GYF; 1.
DR PROSITE; PS50829; GYF; 1.
SQ SEQUENCE 1291 AA; 149173 MW; C0E423D99DF0BC5E CRC64;

Query Match 91.9%; Score 34; DB 2; Length 1291;
Best Local Similarity 85.7%; Pred. No. 3e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
Db 451 ILSPVP 457

RESULT 39
Q857Z6_9CAUD PRELIMINARY; PRT; 111 AA.
AC Q857Z6;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Gp15.
OS Mycobacteriophage CJW1.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=205869;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22592660; PubMed=12705866; DOI=10.1016/S0092-8674(03)00233-2;
RA Pedulla M.L., Ford M.E., Houtz J.M., Karthikeyan T., Wadsworth C.,
RA Lewis J.A., Jacobs-Sera D., Falbo J., Gross J., Pannunzio N.R.,
RA Brucker W., Kumar V., Kandasamy J., Keenan L., Bardarov S.,
RA Kriakov J., Lawrence J.G., Jacobs W.R. Jr., Hendrix R.W.,
RA Hatfull G.F.;
RT "Origins of highly mosaic mycobacteriophage genomes.";
RL Cell 113:171-182 (2003).
DR EMBL; AY129331; AAN01630.1; -; Genomic DNA.
SQ SEQUENCE 111 AA; 12237 MW; AE6B87B381477527 CRC64;

Query Match 89.2%; Score 33; DB 2; Length 111;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAPPVP 7
Db 62 LAPPVP 67

RESULT 40
Y124_LISIN
ID Y124 LISIN STANDARD; PRT; 113 AA.
AC Q925X3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical protein Lin0124/Lin2378/Lin2564.
```

```
GN OrderedLocusNames=lin0124;
GN and
GN OrderedLocusNames=lin2378;
GN and
GN OrderedLocusNames=lin2564;
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CLIP 11262 / Serovar 6a;
RX MEDLINE=21537279; PubMed=11679669; DOI=10.1126/science.1063447;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordbiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852 (2001).
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AL596163; CAC95357.1; -; Genomic DNA.
DR EMBL; AL596172; CAC97605.1; -; Genomic DNA.
DR EMBL; AL596172; CAC97791.1; -; Genomic DNA.
DR PIR; AE1448; AE1448.
DR PIR; AE1729; AE1729.
DR PIR; AG1752; AG1752.
DR Listlist; LIN00124; -.
DR Listlist; LIN02378; -.
DR Listlist; LIN02564; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 113 AA; 12527 MW; D2B7F93D13BD4E32 CRC64;

Query Match 89.2%; Score 33; DB 1; Length 113;
Best Local Similarity 71.4%; Pred. No. 3.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
Db 36 LIAPPVP 42

RESULT 41
Q8DKV4_SYNEL
ID Q8DKV4_SYNEL PRELIMINARY; PRT; 129 AA.
AC Q8DKV4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dihydroneopterin aldolase.
GN OrderedLocusNames=tll0747;
OS Synechococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=32046;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BP-1;
RX MEDLINE=22225144; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shimpo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
```

RP "Complete genome structure of the thermophilic cyanobacterium
RT Thermosynechococcus elongatus BP-1.";
RL DNA Res. 9:123-130(2002).
DR EMBL; BA000039; BAC08298.1; -; Genomic_DNA.
DR HSSP; P56740; 1DHN.
DR GO; GO:0004150; P:dihydroneopterin aldolase activity; IEA.
DR GO; GO:0006760; P:folic acid and derivative metabolism; IEA.
DR InterPro; IPR006157; FolB.
DR InterPro; IPR006156; FolB_fam.
DR Pfam; PF02152; FolB; 1.
DR TIGRFAMs; TIGR00525; folB; 1.
DR TIGRFAMs; TIGR00526; folB_dom; 1.
KW Complete proteome.
SQ SEQUENCE 129 AA; 14567 MW; 689CC2E1C5BECD20 CRC64;

Query Match 89.2%; Score 33; DB 2; Length 129;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAPPVP 7
| | | | |
Db 106 LAPPVP 111

RESULT 42
Q9R6A1_9RHIZ
ID Q9R6A1_9RHIZ PRELIMINARY; PRT; 147 AA.
AC Q9R6A1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Tiorf176 protein.
GN Name=tiorf176;
OS Agrobacterium tumefaciens.
OG Plasmid pTi-SAKURA.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Agrobacterium.
OX NCBI_TaxID=358;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MAFF301001;
RA Hattori Y., Suzuki K., Ohta N., Uraji M., Katoh A., Yoshida K.;
RT "Genome structure of pTi-SAKURA (I): Strategy for DNA sequencing of a
RT Japanese cherry-Ti plasmid."
RL Nucleic Acids Symp. Ser. 37:159-160(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MAFF301001;
RA Ohta N., Suzuki K., Hattori Y., Uraji M., Katoh A., Yoshida K.;
RT "Genome structure of pTi-SAKURA (III): Characteristics of T-DNA."
RL Nucleic Acids Symp. Ser. 39:185-186(1998).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MAFF301001;
RA Uraji M., Suzuki K., Ohta N., Hattori Y., Katoh A., Yoshida K.;
RT "Genome structure of pTi-SAKURA (IV): Characteristics of tra region."
RL Nucleic Acids Symp. Ser. 39:187-188(1998).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MAFF301001;
RA Hattori Y., Suzuki K., Ohta N., Uraji M., Katoh A., Yoshida K.;
RT "Genome structure of pTi-SAKURA (V): Complete nucleotide sequence of
RT plasmid pTi-SAKURA's vir region in Agrobacterium tumefaciens."
RL Nucleic Acids Symp. Ser. 39:265-266(1998).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MAFF301001;
RX MEDLINE=20184752; PubMed=10721727; DOI=10.1016/S0378-1119(99)00502-8;
RA Suzuki K., Hattori Y., Uraji M., Ohta N., Iwata K., Murata K.,
RA Katoh A., Yoshida K.;
RT "Complete nucleotide sequence of a plant tumor-inducing Ti plasmid."
RL Gene 242:331-336(2000).
RN [6]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MAFF301001;
RX MEDLINE=98193120; PubMed=9524202; DOI=10.1016/S0167-4781(97)00182-6;
RA Suzuki K., Ohta N., Hattori Y., Uraji M., Katoh A., Yoshida K.;
RT "Novel structural difference between nopaline- and octopine- type trbJ
RT gene:construction of genetic and physical map and sequencing of
trb/traI and rep gene clusters of a new Ti plasmid pTi-SAKURA.";
RL Biochim. Biophys. Acta 1396:1-7(1998).
DR EMBL; AB016260; BAA87801.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004803; F:transposase activity; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR InterPro; IPR002514; Transposase_8.
DR Pfam; PF01527; Transposase_8; 1.
KW Plasmid.
SQ SEQUENCE 147 AA; 15964 MW; 7B72319E1B9391EE CRC64;

Query Match 89.2%; Score 33; DB 2; Length 147;
Best Local Similarity 71.4%; Pred. No. 4.5e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILAPPVP 7
: | | | | : |
Db 102 VLAPPLP 108

RESULT 43
Q9K3W9_STRCO
ID Q9K3W9_STRCO PRELIMINARY; PRT; 164 AA.
AC Q9K3W9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative marR-family regulatory protein.
GN OrderedLocusNames=SCO4375; ORFNames=SCD10.07;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neil S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939119; CAB95886.1; -; Genomic_DNA.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000835; HTH_MarR.
DR InterPro; IPR011991; Wing_hlx_DNA_bd.
DR Pfam; PF01047; MarR; 1.
DR PRINTS; PR00598; HTHMARR.
DR SMART; SM00347; HTH_MARR; 1.
DR PROSITE; PS50995; HTH_MARR_2; 1.
KW Complete proteome; DNA-binding; Transcription;
KW Transcription regulation.
SQ SEQUENCE 164 AA; 18067 MW; 244067D79F05FEDB CRC64;

Query Match 89.2%; Score 33; DB 2; Length 164;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      2 LAPPVP 7
Db      149 LAPPVP 154

RESULT 44
Q6Z9K3_ORYSA
ID Q6Z9K3_ORYSA PRELIMINARY; PRT; 177 AA.
AC Q6Z9K3;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Hypothetical protein P0475A09.28.
GN Name=P0475A09.28;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 8, PAC
RT clone:P0475A09.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP004696; BAD01331.1; -; Genomic_DNA.
DR Gramene; Q6Z9K3; -.
KW Hypothetical protein.
SQ SEQUENCE 177 AA; 18335 MW; 426AB8BA5E3952BF CRC64;

Query Match      89.2%; Score 33; DB 2; Length 177;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 LAPPVP 7
Db      106 LAPPVP 111

RESULT 45
Q73TA7_MYCPA
ID Q73TA7_MYCPA PRELIMINARY; PRT; 183 AA.
AC Q73TA7;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=MAP3811;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017241; AAS06361.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 183 AA; 16640 MW; D7BE0F54CE6D04E1 CRC64;

Query Match      89.2%; Score 33; DB 2; Length 183;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 LAPPVP 7
Db      118 LAPPVP 123

RESULT 46
HLH4_CAEEL
```

```
ID HLH4_CAEEL STANDARD; PRT; 205 AA.
AC P34555;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Helix-loop-helix protein 4.
GN Name=hlh-4; ORFNames=T05G5.2;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=94150718; PubMed=7906398; DOI=10.1038/368032a0;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA Sims M., Smaldon N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; Z27079; CAA81589.1; -; Genomic_DNA.
CC PIR; S41002; S41002.
CC Ensembl; T05G5.2; Caenorhabditis elegans.
CC WormBase; WBGene00001951; hlh-4.
CC WormPep; T05G5.2; CE00314.
CC InterPro; IPR001092; HLH_basic.
CC Pfam; PF00010; HLH; 1.
CC SMART; SM00353; HLH; 1.
CC PROSITE; PS50888; HLH; 1.
KW Complete proteome; Hypothetical protein; Nuclear protein.
FT DOMAIN 38 89 Helix-loop-helix motif (By similarity).
SQ SEQUENCE 205 AA; 22692 MW; 577E59E04268FA1C CRC64;

Query Match      89.2%; Score 33; DB 1; Length 205;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 LAPPVP 7
Db      139 LAPPVP 144

RESULT 47
P97177_RHOSH
ID P97177_RHOSH PRELIMINARY; PRT; 226 AA.
AC P97177;
DT 01-MAY-1997 (TReMBLrel. 03, Created)
```


DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Transmembrane protein.
GN Name=dmsB;
OS Rhodobacter sphaeroides (Rhodopseudomonas sphaeroides).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC Rhodobacteraceae; Rhodobacter.
OX NCBI_TaxID=1063;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=97107628; PubMed=8950368; DOI=10.1016/S0005-2728(96)00101-6;
RA Ujiye T., Yamamoto I., Nakama H., Okubo A., Yamazaki S., Satoh T.;
RT "Nucleotide sequence of the genes, encoding the pentaheme cytochrome
RT (dmsC) and the transmembrane protein (dmsB), involved in dimethyl
RT sulfoxide respiration from Rhodobacter sphaeroides f. sp.
RT denitrificans.";
RL Biochim. Biophys. Acta 1277:1-5(1996).
DR EMBL; D82820; BAA11582.1; -; Genomic_DNA.
DR PIR; PC4101; PC4101.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0051085; P:chaperone cofactor dependent protein folding; IEA.
DR InterPro; IPR010395; TorD.
DR Pfam; PF06192; TorD; 1.
KW Transmembrane.
SQ SEQUENCE 226 AA; 23551 MW; F821A37FC21E6A3F CRC64;

Query Match 89.2%; Score 33; DB 2; Length 226;
Best Local Similarity 100.0%; Pred. No. 7.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAPPVP 7
Db |||||
3 LAPPVP 8

RESULT 48
Q5FC67 CAEEL
ID Q5FC67 CAEEL PRELIMINARY; PRT; 228 AA.
AC Q5FC67;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein T06E4.12.
GN ORFNames=T06E4.12;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z70756; CAI46587.1; -; Genomic DNA.
DR Ensembl; T06E4.12; Caenorhabditis elegans.
DR WormBase; WBGene00044011; T06E4.12.
DR WormPep; T06E4.12; CE37848.
DR GO; GO:0005199; F:structural constituent of cell wall; IEA.
DR InterPro; IPR003882; Pistil extensin.
DR PRINTS; PR01218; PSTLEXTENSIN.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 228 AA; 22221 MW; FIDEB689484B8913 CRC64;

Query Match 89.2%; Score 33; DB 2; Length 228;
Best Local Similarity 100.0%; Pred. No. 7.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAPPVP 7
Db |||||

Db 112 LAPPVP 117

RESULT 49
Q52WY3 BRARE
ID Q52WY3 BRARE PRELIMINARY; PRT; 235 AA.
AC Q52WY3;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Ras-like protein Rhoub (Fragment).
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15894457; DOI=10.1016/j.ygeno.2005.03.010;
RA Salas-Vidal E., Meijer A.H., Cheng X., Spaink H.P.;
RT "Genomic annotation and expression analysis of the zebrafish Rho small
RT GTPase family during development and bacterial infection.";
RL Genomics 86:25-37(2005).
DR EMBL; AY865565; AAX20137.1; -; mRNA.
DR GO; GO:0005525; F:GTP binding; IEA.
DR GO; GO:0006886; P:intracellular protein transport; IEA.
DR GO; GO:0007264; P:small GTPase mediated signal transduction; IEA.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR003577; GTPase_Ras.
DR InterPro; IPR003578; GTPase_Rho.
DR InterPro; IPR002041; RAN.
DR InterPro; IPR001806; Ras trnsfrmng.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; Ras; 1.
DR PRINTS; PR00449; RASTRNSFRMNG.
DR SMART; SM00175; RAB; 1.
DR SMART; SM00176; RAN; 1.
DR SMART; SM00173; RAS; 1.
DR SMART; SM00174; RHO; 1.
DR TIGRFAMS; TIGR00231; small_GTP; 1.
KW GTP-binding; Nucleotide-binding.
FT NON TER 235 235
SQ SEQUENCE 235 AA; 26491 MW; B79EACC9FC0EF5D4 CRC64;

Query Match 89.2%; Score 33; DB 2; Length 235;
Best Local Similarity 71.4%; Pred. No. 7.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAPPVP 7
Db :|||||
6 LMAPPVP 12

RESULT 50
Q89EI9 BRAJA
ID Q89EI9 BRAJA PRELIMINARY; PRT; 236 AA.
AC Q89EI9;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Bll7096 protein.
GN OrderedLocusNames=bll7096;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,

RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; BA000040; BAC52361.1; -; Genomic_DNA.
DR InterPro; IPR007400; DUF453.
DR Pfam; PF04303; DUF453; 1.
KW Complete proteome.
SQ SEQUENCE 236 AA; 24727 MW; BB79EEB04F48421B CRC64;

Query Match 89.2%; Score 33; DB 2; Length 236;
Best Local Similarity 100.0%; Pred.No. 7.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAPPVP 7
| | | | |
Db 183 LAPPVP 188

Search completed: April 6, 2006, 09:39:26
Job time : 112.105 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2006, 09:22:55 ; Search time 78.2895 Seconds
(without alignments)
39.286 Million cell updates/sec

Title: US-10-632-388-305
Perfect score: 45
Sequence: 1 RQP PPP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	7	3 AAB17249	Aab17249 SH3 antag
2	45	100.0	7	5 ABB73242	Abb73242 Src homol
3	45	100.0	7	7 ADJ73396	Adj73396 SH3 antag
4	45	100.0	7	8 ADJ53030	Adj53030 CH1 delet
5	45	100.0	7	8 ADJ51991	Adj51991 CH1 delet
6	45	100.0	13	2 AAW11113	Aaw11113 Src SH3 d
7	45	100.0	31	2 AAW25511	Aaw25511 Random pe
8	45	100.0	122	8 ADI45407	Adi45407 Rice isop
9	45	100.0	173	5 ABB97921	Abb97921 Human sec
10	45	100.0	249	7 ABO72442	Abo72442 Pseudomon
11	45	100.0	276	8 ADP29798	Adp29798 Human sec
12	45	100.0	288	3 AAG13899	Aag13899 Arabidops
13	45	100.0	314	9 ABM97421	Abm97421 M. xanthu
14	45	100.0	322	3 AAG13898	Aag13898 Arabidops
15	45	100.0	340	3 AAG06371	Aag06371 Arabidops
16	45	100.0	398	3 AAG06370	Aag06370 Arabidops
17	45	100.0	398	8 ADN73283	Adn73283 Thale cre
18	45	100.0	412	3 AAG06369	Aag06369 Arabidops
19	45	100.0	432	2 AAW72182	Aaw72182 HSV-2 str
20	45	100.0	446	4 ABB70063	Abb70063 Drosophil
21	45	100.0	497	2 AAW72030	Aaw72030 HSV-2 str
22	45	100.0	497	2 AAW72132	Aaw72132 HSV-2 str
23	45	100.0	527	2 AAY29515	Aay29515 Human lun
24	45	100.0	527	3 AAB44492	Aab44492 Human lun

25	45	100.0	527	4	AAE13805	Aael3805 Human lun
26	45	100.0	527	7	ADD6524	Add6524 Human lun
27	45	100.0	527	7	ADE87778	Ade87778 Human lun
28	45	100.0	668	5	AAU10545	Aaul0545 Rat synap
29	45	100.0	669	9	ADX05987	Adx05987 Cyclin-de
30	45	100.0	704	5	AAU10544	Aaul0544 Rat synap
31	45	100.0	704	7	ADE57712	Ade57712 Rat Prote
32	45	100.0	705	4	AAM39328	Aam39328 Human pol
33	45	100.0	705	9	ADX05985	Adx05985 Cyclin-de
34	45	100.0	714	4	AAM41114	Aam41114 Human pol
35	45	100.0	735	4	ABG26213	Abg26213 Novel hum
36	45	100.0	812	4	ABG30148	Abg30148 Novel hum
37	45	100.0	874	4	ABG02079	Abg02079 Novel hum
38	45	100.0	1074	4	ABG00401	Abg00401 Novel hum
39	45	100.0	1413	5	AAE21729	Aae21729 Human PKI
40	45	100.0	1460	7	ADC59465	Adc59465 Novel hum
41	45	100.0	1460	8	ADU04632	Adu04632 Human KIA
42	45	100.0	1734	6	ABR42259	Abr42259 Human MAS
43	45	100.0	1734	7	ABM78986	Abm78986 Modifier
44	45	100.0	1734	8	ADG20497	Adg20497 Human MAS
45	45	100.0	1734	8	ADQ15178	Adq15178 Human can
46	45	100.0	1798	5	AAM49177	Aam49177 Human MAS
47	45	100.0	1798	5	AAE16277	Aae16277 Human kin
48	45	100.0	1798	8	ADJ96613	Adj96613 Human mic
49	45	100.0	1805	8	ADX91705	Adx91705 Plant ful
50	45	100.0	4952	5	ADH47759	Adh47759 NOV18 pro
51	45	100.0	4952	7	ADP68294	Adp68294 Human NOV
52	45	100.0	4952	8	ADL25642	Adl25642 Human dia
53	45	100.0	4981	8	ADN05339	Adn05339 Antipsori
54	45	100.0	5159	5	ADH48828	Adh48828 NOVA7 pro
55	42	93.3	129	8	ADX94859	Adx94859 Plant ful
56	42	93.3	216	4	ABB65033	Abb65033 Drosophil
57	42	93.3	279	8	ADR10339	Adr10339 Human pro
58	42	93.3	424	7	ABM88266	Abm88266 Rice abio
59	42	93.3	440	7	ABM90225	Abm90225 Rice abio
60	42	93.3	549	8	ADI45327	Adi45327 Rice isop
61	42	93.3	626	4	AAM40777	Aam40777 Human pol
62	42	93.3	1222	3	AAG36453	Aag36453 Arabidops
63	42	93.3	1257	3	AAG36452	Aag36452 Arabidops
64	42	93.3	1264	8	ADQ66678	Adq66678 Novel hum
65	42	93.3	1275	3	AAG36451	Aag36451 Arabidops
66	42	93.3	1278	2	AAW88445	Aaw88445 Human NPC
67	42	93.3	1278	3	AAB42983	Aab42983 Human ORF
68	42	93.3	1278	8	ADQ39879	Adq39879 Human myo
69	42	93.3	1278	8	ADU06723	Adu06723 Novel bro
70	42	93.3	1639	4	ABB59281	Abb59281 Drosophil
71	42	93.3	1745	5	ABB57253	Abb57253 Mouse isc
72	41	91.1	68	4	ABG11143	Abg11143 Novel hum
73	41	91.1	68	5	AAE26653	Aae26653 Human hun
74	41	91.1	68	5	ABP64470	Abp64470 Human ORF
75	41	91.1	68	6	ABP72613	Abp72613 Huntingto
76	41	91.1	70	8	ADT93238	Adt93238 Human Hun
77	41	91.1	79	4	AAB69616	Aab69616 Huntingto
78	41	91.1	87	5	ABG30880	Abg30880 Human pro
79	41	91.1	87	6	ABP72612	Abp72612 Huntingto
80	41	91.1	89	7	ABM89363	Abm89363 Rice abio
81	41	91.1	89	7	ADU87450	Adu87450 Protein f
82	41	91.1	89	9	ADV77066	Adv77066 Huntingto
83	41	91.1	91	8	ADH09438	Adh09438 Human hun
84	41	91.1	92	8	ADU82410	Adu82410 Wild type
85	41	91.1	92	8	ADU82414	Adu82414 Mutant Hu
86	41	91.1	92	8	ADU82417	Adu82417 Mutant Hu
87	41	91.1	92	8	ADU82419	Adu82419 Mutant Hu
88	41	91.1	92	8	ADU82416	Adu82416 Mutant Hu
89	41	91.1	92	8	ADU82415	Adu82415 Mutant Hu
90	41	91.1	92	8	ADU82418	Adu82418 Mutant Hu
91	41	91.1	92	8	ADU82420	Adu82420 Mutant Hu
92	41	91.1	98	4	AAU62338	Aau62338 Propionib
93	41	91.1	98	6	ABM58857	Abm58857 Propionib
94	41	91.1	108	2	AAW95071	Aaw95071 Amino aci
95	41	91.1	108	2	AAW95076	Aaw95076 Amino aci
96	41	91.1	108	8	ADJ48368	Adj48368 Maize oil
97	41	91.1	112	8	ADX67103	Adx67103 Plant ful

98	41	91.1	114	8	ADS15397	Ads15397 HSV-1 pol	171	41	91.1	724	8	ABO84902	Abo84902 Human can
99	41	91.1	121	9	ADV77067	Adv77067 Huntingto	172	41	91.1	726	7	ABO71929	Abo71929 Pseudomon
100	41	91.1	122	3	AAG33269	Aag33269 Zea mays	173	41	91.1	733	4	ABB71995	Abb71995 Drosophil
101	41	91.1	142	4	ABG18931	Abg18931 Novel hum	174	41	91.1	736	5	ABB57361	Abb57361 Mouse isc
102	41	91.1	142	7	ADF60189	Adf60189 Human con	175	41	91.1	745	8	ABO84899	Abo84899 Murine ca
103	41	91.1	144	7	ABO77746	Abo77746 Pseudomon	176	41	91.1	780	8	ADI82549	Adi82549 Human mod
104	41	91.1	157	7	ABM88176	Abm88176 Rice abio	177	41	91.1	784	3	AAB11694	Aab11694 Human sec
105	41	91.1	164	8	ADU82409	Adu82409 Mutant Un	178	41	91.1	787	3	AAB11693	Aab11693 Human sec
106	41	91.1	171	2	AAW99022	Aaw99022 Human hun	179	41	91.1	825	7	ADJ70479	Adj70479 Human hea
107	41	91.1	171	5	AAE26650	Aae26650 Human hun	180	41	91.1	826	8	ADM48205	Adm48205 Polypepti
108	41	91.1	194	8	ADU01204	Adu01204 Human pro	181	41	91.1	837	3	AA94903	Aay94903 Human sec
109	41	91.1	194	8	ADU15586	Adul15586 Novel hum	182	41	91.1	900	3	AAB42321	Aab42321 Human ORF
110	41	91.1	197	8	ADX68833	Adx68833 Plant ful	183	41	91.1	945	2	AA906119	Aay06119 Human CII
111	41	91.1	216	7	ABM89871	Abm89871 Rice abio	184	41	91.1	972	5	ABB92750	Abb92750 Herbicida
112	41	91.1	218	8	ADX87616	Adx87616 Plant ful	185	41	91.1	978	9	ADU68354	Adu68354 Human klo
113	41	91.1	226	4	ABB71956	Abb71956 Drosophil	186	41	91.1	979	2	AAR08338	Aar08338 Complete
114	41	91.1	237	8	ADY13056	Ady13056 Plant ful	187	41	91.1	1012	2	AAW63671	Aaw63671 Polypepti
115	41	91.1	248	8	ADQ26769	Adq26769 Human her	188	41	91.1	1012	4	AAB73614	Aab73614 Human klo
116	41	91.1	248	8	ADQ26768	Adq26768 Human her	189	41	91.1	1012	7	ADJ69610	Adj69610 Human hea
117	41	91.1	248	8	ADS15399	Ads15399 HSV-1 pol	190	41	91.1	1012	8	ADP43058	Adf43058 Human klo
118	41	91.1	255	6	ABP56850	Abp56850 Zebrafish	191	41	91.1	1012	8	ADQ39809	Adq39809 Human myo
119	41	91.1	264	7	ABM89291	Abm89291 Rice abio	192	41	91.1	1012	9	ADU68353	Adu68353 Human klo
120	41	91.1	309	8	ADY04216	Ady04216 Plant ful	193	41	91.1	1048	4	ABG20939	Abg20939 Novel hum
121	41	91.1	315	9	ABE09208	Aeb09208 Huntingti	194	41	91.1	1048	8	ADS11844	Ads11844 Human the
122	41	91.1	317	4	AAU67582	Aau67582 Propionib	195	41	91.1	1109	7	ADC59312	Adc59312 Human pol
123	41	91.1	317	6	ABM64101	Abm64101 Propionib	196	41	91.1	1217	4	ABG09876	Abg09876 Novel hum
124	41	91.1	322	9	ADV77064	Adv77064 Huntingto	197	41	91.1	1239	4	ABG09877	Abg09877 Novel hum
125	41	91.1	330	7	ABO69642	Abo69642 Pseudomon	198	41	91.1	1277	9	AEB32343	Aeb32343 Human pro
126	41	91.1	331	8	ADT93240	Adt93240 Human Hun	199	41	91.1	1278	7	ABM85431	Abm85431 Human pro
127	41	91.1	334	6	ABM64819	Abm64819 Propionib	200	41	91.1	1313	7	ADJ70374	Adj70374 Human hea
128	41	91.1	334	7	ADM05556	Adm05556 Human pro	201	41	91.1	1324	8	ADJ66660	Adj66660 IRS2 prot
129	41	91.1	357	7	ADJ80172	Adj80172 Novel hum	202	41	91.1	1324	8	ADN98330	Adn98330 Human ins
130	41	91.1	361	8	ADI42062	Adi42062 Plant tra	203	41	91.1	1324	8	ADR14675	Adr14675 Human NF-
131	41	91.1	361	8	ADO02607	Ado02607 Thalecres	204	41	91.1	1324	9	ADU87428	Adu87428 Human IRS
132	41	91.1	380	8	ADY24834	Ady24834 Plant ful	205	41	91.1	1324	9	AEB32315	Aeb32315 Human pro
133	41	91.1	384	8	ADY04877	Ady04877 Plant ful	206	41	91.1	1338	8	ADO19685	Ado19685 Human PRO
134	41	91.1	387	6	ABO00597	Abo00597 Novel hum	207	41	91.1	1338	8	ADP54811	Adp54811 Human PRO
135	41	91.1	406	6	ABP58350	Abp58350 Human cel	208	41	91.1	1338	8	ADP24219	Adp24219 PRO polyp
136	41	91.1	413	9	ADW17172	Adw17172 Eucalyptu	209	41	91.1	1338	8	ADT07511	Adt07511 Human col
137	41	91.1	422	3	AAG41764	Aag41764 Arabidops	210	41	91.1	1340	6	AAE37017	Aae37017 Human nuc
138	41	91.1	440	4	AAU30409	Aau30409 Novel hum	211	41	91.1	1400	9	ADV85501	Adv85501 Mouse bro
139	41	91.1	465	8	ADT93239	Adt93239 Human Hun	212	41	91.1	1420	4	ABB63410	Abb63410 Drosophil
140	41	91.1	475	7	ABM87048	Abm87048 Rice abio	213	41	91.1	1430	7	ADJ69277	Adj69277 Human hea
141	41	91.1	481	8	ADT02468	Adt02468 Human CRE	214	41	91.1	1506	9	ADZ13633	Adz13633 Human can
142	41	91.1	485	4	ABE58326	Abb58326 Drosophil	215	41	91.1	1506	9	ADZ13622	Adz13622 Human can
143	41	91.1	485	6	ABP58349	Abp58349 Human cel	216	41	91.1	1509	8	ADT07512	Adt07512 Human col
144	41	91.1	485	7	ADC31623	Adc31623 Human nov	217	41	91.1	1509	8	ADT07513	Adt07513 Human col
145	41	91.1	485	7	ADM05364	Adm05364 Human pro	218	41	91.1	1542	5	ABB78013	Abb78013 Amino aci
146	41	91.1	489	8	ADY10556	Ady10556 Plant ful	219	41	91.1	1616	8	ADP22958	Adp22958 PRO polyp
147	41	91.1	494	8	ADT02466	Adt02466 Human CRE	220	41	91.1	1638	7	ADC31236	Adc31236 Human nov
148	41	91.1	502	8	ADP12999	Adp12999 Protein e	221	41	91.1	1682	8	ADR14131	Adr14131 Human NF-
149	41	91.1	513	2	AA933500	Aay33500 Human hun	222	41	91.1	2000	8	ADP23488	Adp23488 PRO polyp
150	41	91.1	516	9	ADV77063	Adv77063 Huntingto	223	41	91.1	2000	9	ADY14779	Ady14779 PRO polyp
151	41	91.1	530	2	AA933501	Aay33501 Human apo	224	41	91.1	2004	5	ABG95113	Abg95113 Human tra
152	41	91.1	535	7	ABM89001	Abm89001 Rice abio	225	41	91.1	2004	9	ADX07185	Adx07185 Cyclin-de
153	41	91.1	549	2	AAW63672	Aaw63672 Polypepti	226	41	91.1	2004	9	ADY17185	Ady17185 PRO polyp
154	41	91.1	549	4	AAAB73615	Aab73615 Human klo	227	41	91.1	2397	8	ADU02443	Adu02443 Novel hum
155	41	91.1	549	8	ADQ39810	Adq39810 Human myo	228	41	91.1	2614	9	ADZ13635	Adz13635 Human can
156	41	91.1	552	2	AA933502	Aay33502 Human apo	229	41	91.1	2614	9	ADZ13624	Adz13624 Human can
157	41	91.1	568	3	AAAG41763	Aag41763 Arabidops	230	41	91.1	2910	8	ADQ59382	Adq59382 Human can
158	41	91.1	580	8	ADT02467	Adt02467 Human CRE	231	41	91.1	2910	9	ADZ13630	Adz13630 Human can
159	41	91.1	589	2	AA933503	Aay33503 Human apo	232	41	91.1	2910	9	ADZ13641	Adz13641 Human can
160	41	91.1	590	3	AAAG41762	Aag41762 Arabidops	233	41	91.1	2911	6	ABU56630	Abu56630 Lung canc
161	41	91.1	593	4	AAAB95646	Aab95646 Human pro	234	41	91.1	2911	6	ABO07259	Abo07259 Human p53
162	41	91.1	593	6	AAAG79911	Aag79911 MECT1. 4/	235	41	91.1	2911	7	ADJ68615	Adj68615 Human hea
163	41	91.1	597	8	ADU02790	Adu02790 Novel hum	236	41	91.1	2911	7	ADN38844	Adn38844 Cancer/an
164	41	91.1	629	8	ADT02461	Adt02461 mCREAP1.	237	41	91.1	2911	8	ADQ18204	Adq18204 Human sof
165	41	91.1	634	8	ABO84900	Abo84900 Human can	238	41	91.1	2911	8	ADU06413	Adu06413 Novel bro
166	41	91.1	650	4	AAU28167	Aau28167 Novel hum	239	41	91.1	2912	4	ABG06402	Abg06402 Novel hum
167	41	91.1	650	8	ADT02434	Adt02434 Human CRE	240	41	91.1	2912	9	ADZ13626	Adz13626 Human can
168	41	91.1	657	4	AAU28355	Aau28355 Novel hum	241	41	91.1	2912	9	ADZ13637	Adz13637 Human can
169	41	91.1	716	2	AAW38192	Aaw38192 Maize ZCA	242	41	91.1	2912	9	ADZ13628	Adz13628 Human can
170	41	91.1	716	3	AAAB28591	Aab28591 Maize ZCR	243	41	91.1	2912	9	ADZ13639	Adz13639 Human can

244	41	91.1	3051	6	ABR58308	Abr58308	BCU0947 p	317	40	88.9	77	4	AAO02698	Aao02698	Human pol
245	41	91.1	3124	8	ADJ78494	Adj78494	E1A bindi	318	40	88.9	78	4	AAM89336	Aam89336	Human imm
246	41	91.1	3141	7	ADJ70444	Adj70444	Human hea	319	40	88.9	78	4	AAO03851	Aao03851	Human pol
247	41	91.1	3144	2	AAR58777	Aar58777	Protein e	320	40	88.9	80	5	AAM49170	Aam49170	Human tra
248	41	91.1	3144	2	AAW36887	Aaw36887	Previousel	321	40	88.9	82	3	AAB51988	Aab51988	Human sec
249	41	91.1	3144	2	AAW09871	Aaw09871	Human hun	322	40	88.9	92	4	AAU32238	Aau32238	Novel hum
250	41	91.1	3144	2	AAW44742	Aaw44742	Human hun	323	40	88.9	95	4	ABH11356	Abh11356	Human sec
251	41	91.1	3144	2	AAW33493	Aay33493	Human hun	324	40	88.9	95	4	AAO07303	Aao07303	Human pol
252	41	91.1	3144	9	ADY98141	Ady98141	Human hun	325	40	88.9	95	5	ABP60103	Abp60103	DNA bindi
253	41	91.1	3223	4	ABBI1407	Abb11407	Human Hun	326	40	88.9	96	3	AAG25330	Agg25330	Zea may
254	41	91.1	3223	4	ABBI1470	Abb11470	Human Hun	327	40	88.9	97	8	ADM87815	Adm87815	Human EST
255	41	91.1	3469	8	ADS96634	Ads96634	Drosophil	328	40	88.9	98	4	ABG10944	Abg10944	Novel hum
256	41	91.1	3502	4	ABB58382	Abb58382	Drosophil	329	40	88.9	98	5	ABJ01042	Abj01042	Human bre
257	41	91.1	3604	8	ADQ89594	Adq89594	Antagonis	330	40	88.9	98	5	ADG79629	Adg79629	Human sec
258	40	88.9	9	6	ABR28252	Abr28252	Human can	331	40	88.9	99	5	ABP08829	Abp08829	Human ORF
259	40	88.9	10	6	ABR28347	Abr28347	Human can	332	40	88.9	100	4	ABG25413	Abg25413	Novel hum
260	40	88.9	15	2	AAW39016	Aaw39016	Peptide r	333	40	88.9	106	4	ABG07829	Abg07829	Novel hum
261	40	88.9	15	2	AAW38985	Aaw38985	Peptide r	334	40	88.9	107	4	AAO0808	Aao0808	Human pol
262	40	88.9	15	6	ABR38294	Abr38294	Human can	335	40	88.9	107	6	ADA55418	Ada55418	Human pro
263	40	88.9	15	6	ABR38391	Abr38391	Human can	336	40	88.9	108	4	ABB69539	Abb69539	Drosophil
264	40	88.9	15	6	ABR38293	Abr38293	Human can	337	40	88.9	108	5	ABG62148	Abg62148	Human pro
265	40	88.9	15	6	ABR38392	Abr38392	Human can	338	40	88.9	109	4	AAO01024	Aao01024	Human pol
266	40	88.9	15	6	ABR38351	Abr38351	Human can	339	40	88.9	110	4	ABB68326	Abb68326	Drosophil
267	40	88.9	15	6	ABR38292	Abr38292	Human can	340	40	88.9	110	4	AAB95822	Aab95822	Human pro
268	40	88.9	30	4	AAO02051	Aao02051	Human pol	341	40	88.9	112	4	AAO02189	Aao02189	Human pol
269	40	88.9	33	4	AAO12108	Aao12108	Human pol	342	40	88.9	113	5	ABP57768	Abp57768	Human zin
270	40	88.9	37	6	AAO19948	Aao19948	Alzheimer	343	40	88.9	114	4	AAO05912	Aao05912	Human pol
271	40	88.9	38	6	AAO19947	Aao19947	Alzheimer	344	40	88.9	114	4	AAO02218	Aao02218	Human pol
272	40	88.9	41	1	AAO10196	Aap10196	Sequence	345	40	88.9	114	7	ADC33200	Adc33200	Human nov
273	40	88.9	41	4	AAO05676	Aao05676	Human pol	346	40	88.9	115	4	AAU32525	Aau32525	Novel hum
274	40	88.9	42	1	AAP71006	Aap71006	Sequence	347	40	88.9	115	4	AAU32143	Aau32143	Novel hum
275	40	88.9	42	1	AAP80040	Aap80040	Beta-huma	348	40	88.9	115	5	ABR01725	Abr01725	Human bre
276	40	88.9	42	1	AAP91846	Aap91846	Anaologue	349	40	88.9	116	4	ABB66710	Abb66710	Drosophil
277	40	88.9	42	2	AAW42213	Aaw42213	Peptide d	350	40	88.9	116	4	AAO00466	Aao00466	Human pol
278	40	88.9	42	2	AAW69459	Aaw69459	HCG antig	351	40	88.9	116	5	ADG79555	Adg79555	Human sec
279	40	88.9	42	2	AAW93444	Aaw93444	Human hCG	352	40	88.9	116	5	ADG79456	Adg79456	Human sec
280	40	88.9	42	3	AAW87489	Aay87489	Human cho	353	40	88.9	116	8	ADQ66101	Adq66101	Novel hum
281	40	88.9	42	3	AAW20553	Aab20553	Human cho	354	40	88.9	117	7	ADD19326	Add19326	Human sec
282	40	88.9	42	4	AAU01149	Aau01149	Structure	355	40	88.9	117	7	ADD19292	Add19292	Human sec
283	40	88.9	42	4	AAW48395	Aab48395	Human cho	356	40	88.9	118	7	ADB37548	Adb37548	Neural th
284	40	88.9	42	4	AAU02847	Aau02847	Human Cho	357	40	88.9	120	8	ADX70886	Adx70886	Plant ful
285	40	88.9	42	4	AAW04131	Aab04131	Peptide f	358	40	88.9	121	4	AAO13020	Aao13020	Human pol
286	40	88.9	43	5	AAW50788	Aam50788	Human cho	359	40	88.9	121	4	ABG16589	Abg16589	Novel hum
287	40	88.9	46	4	AAO10556	Aao10556	Human pol	360	40	88.9	125	4	AAO04869	Aao04869	Human pol
288	40	88.9	46	4	ABG17583	Abg17583	Novel hum	361	40	88.9	125	7	ADD19266	Add19266	Human sec
289	40	88.9	48	1	AAP10195	Aap10195	Sequence	362	40	88.9	126	4	AAO00766	Aao00766	Human pol
290	40	88.9	48	1	AAP80039	Aap80039	Beta-huma	363	40	88.9	130	8	ADR08968	Adr08968	Human pro
291	40	88.9	48	1	AAP91845	Aap91845	Anaologue	364	40	88.9	131	4	AAO01747	Aao01747	Human pol
292	40	88.9	48	2	AAW42212	Aaw42212	Peptide d	365	40	88.9	132	4	AAO00570	Aao00570	Human pol
293	40	88.9	48	2	AAW69458	Aaw69458	HCG antig	366	40	88.9	132	4	AAO03924	Aao03924	Human pol
294	40	88.9	48	2	AAW93443	Aaw93443	Human hCG	367	40	88.9	135	7	ADE08953	Ade08953	Novel pro
295	40	88.9	48	3	AAW87488	Aay87488	Human cho	368	40	88.9	135	8	ADS11754	Ads11754	Human the
296	40	88.9	48	3	AAW20552	Aab20552	Human cho	369	40	88.9	136	4	AAO01204	Aao01204	Human pol
297	40	88.9	48	4	AAU01148	Aau01148	Structure	370	40	88.9	137	4	ABG12622	Abg12622	Novel hum
298	40	88.9	48	4	AAW48394	Aab48394	Human cho	371	40	88.9	138	8	ADQ66021	Adq66021	Novel hum
299	40	88.9	48	4	AAU02846	Aau02846	Human Cho	372	40	88.9	141	4	AAO05999	Aao05999	Human pol
300	40	88.9	48	4	AAW04130	Aab04130	Peptide f	373	40	88.9	144	5	ADK37055	Adk37055	Novel hum
301	40	88.9	48	5	ADK36186	Adk36186	Novel hum	374	40	88.9	144	5	ADK37004	Adk37004	Novel hum
302	40	88.9	48	5	ADK35869	Adk35869	Novel hum	375	40	88.9	146	6	ABP75894	Abp75894	Human sec
303	40	88.9	50	4	AAU49050	Aau49050	Propionib	376	40	88.9	148	8	ADQ65118	Adq65118	Novel hum
304	40	88.9	50	6	ABM45569	Abm45569	Propionib	377	40	88.9	148	8	ADQ66558	Adq66558	Novel hum
305	40	88.9	52	4	AAO11780	Aao11780	Human pol	378	40	88.9	149	4	AAO03286	Aao03286	Human pol
306	40	88.9	53	2	AAW26916	Aay26916	MLV MCF e	379	40	88.9	152	4	ABG16591	Abg16591	Novel hum
307	40	88.9	56	4	AAO05154	Aao05154	Human pol	380	40	88.9	152	8	ADQ66482	Adq66482	Novel hum
308	40	88.9	56	4	ABG12672	Abg12672	Novel hum	381	40	88.9	154	4	ABG08425	Abg08425	Novel hum
309	40	88.9	57	4	AAO04578	Aao04578	Human pol	382	40	88.9	155	5	AAU72798	Aau72798	Human ant
310	40	88.9	59	4	AAO05645	Aao05645	Human pol	383	40	88.9	156	4	ABG10943	Abg10943	Novel hum
311	40	88.9	63	4	AAO10207	Aao10207	Human pol	384	40	88.9	158	6	ADA83852	Ada83852	Human POM
312	40	88.9	64	5	ABP05328	Abp05328	Human ORF	385	40	88.9	164	4	AAU27698	Aau27698	Full-leng
313	40	88.9	73	7	ABM87180	Abm87180	Rice abio	386	40	88.9	172	4	ABG19542	Abg19542	Novel hum
314	40	88.9	75	4	AAW69169	Aab69169	Rat VGF p	387	40	88.9	174	3	AAB43208	Aab43208	Human ORF
315	40	88.9	75	4	AAW69547	Aab69547	Rat VGF8	388	40	88.9	175	4	AAU21256	Aau21256	Human nov
316	40	88.9	77	4	AAO01134	Aao01134	Human pol	389	40	88.9	180	7	ADB64804	Adb64804	Human pro

390	40	88.9	184	8	ADI42026	Adi42026 Plant tra	463	40	88.9	265	7	ADB85017	Adb85017 Human PRO
391	40	88.9	184	8	ADO02576	Ado02576 Thalecres	464	40	88.9	265	7	ADB78123	Adb78123 Novel hum
392	40	88.9	184	8	ADO62215	Ado62215 Transcrip	465	40	88.9	265	7	ADB87189	Adb87189 Human PRO
393	40	88.9	197	4	ABG22105	Abg22105 Novel hum	466	40	88.9	265	7	ADB84771	Adb84771 Human PRO
394	40	88.9	201	4	ABG19534	Abg19534 Novel hum	467	40	88.9	265	7	ADB83886	Adb83886 Novel hum
395	40	88.9	207	4	AAU27870	Aau27870 Contig po	468	40	88.9	265	7	ADB73041	Adb73041 Novel hum
396	40	88.9	217	8	ADI42771	Adi42771 Plant tra	469	40	88.9	265	7	ADC57703	Adc57703 Human PRO
397	40	88.9	218	7	ADC86939	Adc86939 Human GPC	470	40	88.9	265	7	ADC55067	Adc55067 Human PRO
398	40	88.9	219	4	ACG66500	Aag66500 Human neu	471	40	88.9	265	7	ADC11934	Adc11934 Human sec
399	40	88.9	219	7	ABM85689	Abm85689 Human pro	472	40	88.9	265	7	ADC56356	Adc56356 Human PRO
400	40	88.9	220	4	AAU30601	Aau30601 Novel hum	473	40	88.9	265	7	ADC07411	Adc07411 Human sec
401	40	88.9	228	4	AAB60487	Aab60487 Human cel	474	40	88.9	265	7	ADC11401	Adc11401 Human sec
402	40	88.9	230	5	ABP43974	Abp43974 FLJ23018	475	40	88.9	265	7	ADC36879	Adc36879 Human PRO
403	40	88.9	232	4	AAW24472	Aam24472 Human EST	476	40	88.9	265	7	ADC21869	Adc21869 Human PRO
404	40	88.9	232	7	ADI60378	Adi60378 Secreted	477	40	88.9	265	7	ADC49900	Adc49900 Novel hum
405	40	88.9	234	4	ABG24732	Abg24732 Novel hum	478	40	88.9	265	7	ADC49099	Adc49099 Novel hum
406	40	88.9	235	6	ABJ19682	Abj19682 Human sec	479	40	88.9	265	7	ADC49616	Adc49616 Novel hum
407	40	88.9	235	6	ABP99572	Abp99572 Human sec	480	40	88.9	265	7	ADC47477	Adc47477 Novel hum
408	40	88.9	235	8	ADT59093	Adt59093 Plant pol	481	40	88.9	265	7	ADC14523	Adc14523 Novel hum
409	40	88.9	236	3	AAB39216	Aab39216 Human sec	482	40	88.9	265	7	ADC47222	Adc47222 Novel hum
410	40	88.9	238	7	ABM89261	Abm89261 Rice abio	483	40	88.9	265	7	ADD08055	Add08055 Novel hum
411	40	88.9	245	8	ADT59408	Adt59408 Plant pol	484	40	88.9	265	7	ADC81880	Adc81880 Human PRO
412	40	88.9	247	4	ABG15898	Abg15898 Novel hum	485	40	88.9	265	7	ADD07522	Add07522 Novel hum
413	40	88.9	247	4	AAG89176	Aag89176 Human sec	486	40	88.9	265	7	ADC78097	Adc78097 Novel hum
414	40	88.9	252	8	ADX96980	Adx96980 Plant ful	487	40	88.9	265	7	ADC82413	Adc82413 Human PRO
415	40	88.9	261	7	ABM89090	Abm89090 Rice abio	488	40	88.9	265	7	ADD06332	Add06332 Novel hum
416	40	88.9	262	3	AYO8954	Aay08954 Plant ful	489	40	88.9	265	7	ADD08593	Add08593 Novel hum
417	40	88.9	265	8	AAY66691	Aay66691 Membrane-	490	40	88.9	265	7	ADC77851	Adc77851 Novel hum
418	40	88.9	265	3	AAE24063	Aab24063 Human PRO	491	40	88.9	265	7	ADD06842	Add06842 Novel hum
419	40	88.9	265	4	AAB52114	Aab65214 Human PRO	492	40	88.9	265	7	ADC83089	Adc83089 Human PRO
420	40	88.9	265	5	AAU83666	Aau83666 Human PRO	493	40	88.9	265	7	ADD50814	Add50814 Novel hum
421	40	88.9	265	5	ADY31890	Ady31890 Novel hum	494	40	88.9	265	7	ADD51060	Add51060 Novel hum
422	40	88.9	265	6	ABU58029	Abu58029 Human PRO	495	40	88.9	265	7	ADD55196	Add55196 Human PRO
423	40	88.9	265	6	ABU59107	Abu59107 Novel hum	496	40	88.9	265	7	ADD56154	Add56154 Human PRO
424	40	88.9	265	6	ABU82619	Abu82619 Human sec	497	40	88.9	265	7	ADD50541	Add50541 Human PRO
425	40	88.9	265	6	ABU60538	Abu60538 Human sec	498	40	88.9	265	7	ADD54592	Add54592 Human PRO
426	40	88.9	265	6	ABU50813	Abu50813 Human PRO	499	40	88.9	265	7	ADD50295	Add50295 Human PRO
427	40	88.9	265	6	ABO33779	Abu33779 Novel hum	500	40	88.9	265	7	ADD51306	Add51306 Novel hum
428	40	88.9	265	6	ABU13920	Abu13920 Human PRO	501	40	88.9	265	7	ADE26746	Ade26746 Novel hum
429	40	88.9	265	6	ABU72505	Abu72505 Novel hum	502	40	88.9	265	7	ADE26213	Ade26213 Novel hum
430	40	88.9	265	6	ABU59254	Abu59254 Human sec	503	40	88.9	265	7	ADP67150	Adf67150 Human PRO
431	40	88.9	265	6	ABO25951	Abu25951 Human PRO	504	40	88.9	265	7	ADI35404	Adi35404 Human PRO
432	40	88.9	265	6	ABU82122	Abu82122 Novel hum	505	40	88.9	265	7	ADH99896	Adh99896 Novel hum
433	40	88.9	265	6	ABU58960	Abu58960 Human sec	506	40	88.9	265	8	ADC48853	Adc48853 Novel hum
434	40	88.9	265	6	ABU92338	Abu92338 Novel hum	507	40	88.9	265	8	ADE21024	Ade21024 Novel hum
435	40	88.9	265	6	ABU59403	Abu59403 Novel hum	508	40	88.9	265	8	ADE05868	Ade05868 Human PRO
436	40	88.9	265	6	ABU92169	Abu92169 Novel hum	509	40	88.9	265	8	ADD75097	Add75097 Human PRO
437	40	88.9	265	6	ABU10875	Abu10875 Human PRO	510	40	88.9	265	8	ADD75843	Add75843 Novel hum
438	40	88.9	265	6	ABU81627	Abu81627 Novel hum	511	40	88.9	265	8	ADD85075	Add85075 Novel hum
439	40	88.9	265	6	ABU88566	Abu88566 Human sec	512	40	88.9	265	8	ADD86901	Add86901 Novel hum
440	40	88.9	265	6	ABO34080	Abu34080 Human PRO	513	40	88.9	265	8	ADP20778	Ade20778 Novel hum
441	40	88.9	265	6	ABJ72302	Abj72302 Human PRO	514	40	88.9	265	8	ADE39075	Ade39075 Novel hum
442	40	88.9	265	6	ADA37734	Ada37734 Human sec	515	40	88.9	265	8	ADE05622	Ade05622 Human PRO
443	40	88.9	265	6	ADA21420	Ada21420 Human sec	516	40	88.9	265	8	ADD73607	Add73607 Human PRO
444	40	88.9	265	6	ADA10207	Ada10207 Human sec	517	40	88.9	265	8	ADD78447	Add78447 Novel hum
445	40	88.9	265	6	ADA17751	Ada17751 Human PRO	518	40	88.9	265	8	ADE21270	Ade21270 Novel hum
446	40	88.9	265	6	ADA27859	Ada27859 Human sec	519	40	88.9	265	8	ADD77385	Add77385 Novel hum
447	40	88.9	265	6	ADA94439	Ada94439 Human sec	520	40	88.9	265	8	ADE20532	Ade20532 Novel hum
448	40	88.9	265	6	ADA38664	Ada38664 Human sec	521	40	88.9	265	8	ADD75597	Add75597 Human PRO
449	40	88.9	265	6	ABJ72430	Abj72430 Human PRO	522	40	88.9	265	8	ADD74113	Add74113 Human PRO
450	40	88.9	265	6	ADA92785	Ada92785 Human sec	523	40	88.9	265	8	ADD74359	Add74359 Human PRO
451	40	88.9	265	6	ABO34325	Abu34325 Human sec	524	40	88.9	265	8	ADD76089	Add76089 Novel hum
452	40	88.9	265	7	ABO53166	Abu53166 Human sec	525	40	88.9	265	8	ADD85581	Add85581 Novel hum
453	40	88.9	265	7	ADA22346	Ada22346 Human sec	526	40	88.9	265	8	ADE05130	Ade05130 Human PRO
454	40	88.9	265	7	ABO22536	Abu22536 Human sec	527	40	88.9	265	8	ADD75343	Add75343 Human PRO
455	40	88.9	265	7	ADA06512	Ada06512 Human sec	528	40	88.9	265	8	ADD76887	Add76887 Novel hum
456	40	88.9	265	7	ABJ72132	Abj72132 Human mem	529	40	88.9	265	8	ADD86655	Add86655 Novel hum
457	40	88.9	265	7	ADA39205	Ada39205 Human sec	530	40	88.9	265	8	ADD78123	Add78123 Novel hum
458	40	88.9	265	7	ADB83640	Adb83640 Novel hum	531	40	88.9	265	8	ADD77631	Add77631 Novel hum
459	40	88.9	265	7	ADB80746	Adb80746 Novel hum	532	40	88.9	265	8	ADD77877	Add77877 Novel hum
460	40	88.9	265	7	ADB73287	Adb73287 Novel hum	533	40	88.9	265	8	ADD85335	Add85335 Novel hum
461	40	88.9	265	7	ADB96231	Adb96231 Human PRO	534	40	88.9	265	8	ADD73867	Add73867 Human PRO
462	40	88.9	265	7	ADB78369	Adb78369 Novel hum	535	40	88.9	265	8	ADD74605	Add74605 Human PRO

536 40 88.9 265 8 ADD77133 Novel hum
537 40 88.9 265 8 ADD85827 Novel hum
538 40 88.9 265 8 ADE05376 Human PRO
539 40 88.9 265 8 ADD74851 Human PRO
540 40 88.9 265 8 ADF35349 Human PRO
541 40 88.9 265 8 ADG11599 Human PRO
542 40 88.9 265 8 ADG05663 Novel hum
543 40 88.9 265 8 ADG27217 Human PRO
544 40 88.9 265 8 ADG11280 Novel hum
545 40 88.9 265 8 ADG12059 Novel hum
546 40 88.9 265 8 ADF94616 Novel hum
547 40 88.9 265 8 ADG06712 Human PRO
548 40 88.9 265 8 ADH39056 Novel hum
549 40 88.9 265 8 ADH19469 Human sec
550 40 88.9 265 8 ADH20962 Human sec
551 40 88.9 265 8 ADH20002 Human sec
552 40 88.9 265 8 ADG34146 Novel hum
553 40 88.9 265 8 ADI33616 Human PRO
554 40 88.9 265 8 ADH69710 Human PRO
555 40 88.9 265 8 ADI29871 Novel hum
556 40 88.9 265 8 ADM27268 Novel hum
557 40 88.9 265 8 ADK66626 Human PRO
558 40 88.9 265 9 AEA38486 Human sec
559 40 88.9 266 5 AAM48706 Human tra
560 40 88.9 267 8 ADY11763 Plant ful
561 40 88.9 268 8 ADO02317 Thalecres
562 40 88.9 270 8 ADY09279 Plant ful
563 40 88.9 274 8 ADQ66448 Novel hum
564 40 88.9 285 7 ADI60139 Secreted
565 40 88.9 291 7 ABM88444 Rice abio
566 40 88.9 309 7 ADC31415 Human nov
567 40 88.9 310 7 ABM74160 DNA clone
568 40 88.9 311 4 AAM24509 C883P pre
569 40 88.9 311 5 AAB47945 C883P pre
570 40 88.9 311 6 ABP55358 Human col
571 40 88.9 313 8 ADQ89848 Antagonis
572 40 88.9 316 5 AAE22089 Human nov
573 40 88.9 319 3 AAG46613 Arabidops
574 40 88.9 320 3 AAG37027 Arabidops
575 40 88.9 320 8 ADR09376 Human pro
576 40 88.9 322 7 ABM86610 Rice abio
577 40 88.9 323 8 ADI43438 Plant tra
578 40 88.9 329 8 ADR80067 Plant ful
579 40 88.9 332 7 ABM88621 Rice abio
580 40 88.9 333 8 ADR08903 Human pro
581 40 88.9 334 5 ABJ04705 Mycobacte
582 40 88.9 335 7 ABM86681 Rice abio
583 40 88.9 339 7 ABM86928 Rice abio
584 40 88.9 343 3 AAB41666 Human ORF
585 40 88.9 345 5 AAU84271 Human end
586 40 88.9 346 2 AAW14442 Human tra
587 40 88.9 346 7 AAD46201 Human PRO
588 40 88.9 346 7 ADE62409 Human PRO
589 40 88.9 346 9 ADX05441 Cyclin-de
590 40 88.9 350 7 ADG42159 Human bra
591 40 88.9 350 7 ADF74180 Human nov
592 40 88.9 350 9 AEA50208 HNF3G. 8/
593 40 88.9 352 7 ADM04873 Human pro
594 40 88.9 353 5 AAE22090 Human nov
595 40 88.9 354 4 AAB60375 Human ato
596 40 88.9 354 4 AAB60349 Human ato
597 40 88.9 354 6 AAO30945 Human ato
598 40 88.9 354 6 AAO30920 Human ato
599 40 88.9 354 8 ADQ19897 Human sof
600 40 88.9 354 8 ADR69062 Human ato
601 40 88.9 354 8 ADR69063 Human ato
602 40 88.9 356 4 AAB70701 Human TAA
603 40 88.9 357 5 AAU85555 Clone #50
604 40 88.9 357 5 AAU69431 Lung smal
605 40 88.9 357 6 ABU69527 Human lun
606 40 88.9 357 6 ABU66430 Lung canc
607 40 88.9 357 7 ADH47336 Human lun
608 40 88.9 357 8 ADJ21255 Human lun

609 40 88.9 359 9 ADV68459
610 40 88.9 359 9 ADY54996
611 40 88.9 361 7 ABM90394
612 40 88.9 362 7 ABM87638
613 40 88.9 363 4 AAM25490
614 40 88.9 371 7 ABO43147 A. thalia
615 40 88.9 371 8 ADO02519 Thalecres
616 40 88.9 373 4 AAU03507 Human pro
617 40 88.9 376 4 ABB64723 Drosophil
618 40 88.9 389 6 ABU33914 Protein e
619 40 88.9 390 2 AAY17066 Human 3-O
620 40 88.9 390 4 AAB95507 Human pro
621 40 88.9 390 8 ADI09937 Human hep
622 40 88.9 390 8 ADJ76166 Marker ge
623 40 88.9 390 8 ADL61251 Human hep
624 40 88.9 390 9 ADX26316 Novel cel
625 40 88.9 391 4 AAB70114 Murine 3 -
626 40 88.9 391 4 AAB70115 Human 3-O
627 40 88.9 399 3 AAB08218 A protein
628 40 88.9 402 9 ADX83401 Human TEG
629 40 88.9 404 8 ADT58544 Plant pol
630 40 88.9 412 6 ABR42686 Petunia a
631 40 88.9 414 6 AAE32027 Human kin
632 40 88.9 414 8 ADJ96641 Human Nim
633 40 88.9 416 5 AAE15660 Human hom
634 40 88.9 416 6 ABP97648 Amino aci
635 40 88.9 416 9 AEB48139 Amino aci
636 40 88.9 431 3 AAG32324 Arabidops
637 40 88.9 431 3 AAG09613 Arabidops
638 40 88.9 431 8 ADI43535 Plant tra
639 40 88.9 431 9 AEA26765 Stress to
640 40 88.9 432 5 ABP69406 Human pol
641 40 88.9 432 5 ABU65180 Human NOV
642 40 88.9 432 8 ADN62011 Human nov
643 40 88.9 432 9 ADY18827 PRO polyp
644 40 88.9 434 9 ABM91399 M. xanthu
645 40 88.9 435 3 AAG09612 Arabidops
646 40 88.9 439 7 ADF74169 Human nov
647 40 88.9 439 8 ADY25088 Plant ful
648 40 88.9 440 3 AAG49463 Arabidops
649 40 88.9 441 6 ABP75500 Human sec
650 40 88.9 443 3 AAG49462 Arabidops
651 40 88.9 453 7 ADD48354 Human Pro
652 40 88.9 455 7 ABO69519 Pseudomon
653 40 88.9 456 7 ADE40443 Human cho
654 40 88.9 461 4 AAU17200 Novel sig
655 40 88.9 461 7 ADB93908 Human nov
656 40 88.9 461 7 ADJ70999 Human hea
657 40 88.9 461 8 ADY08284 Plant ful
658 40 88.9 463 3 AAB41079 Human ORF
659 40 88.9 471 8 ADJ71959 Human PMM
660 40 88.9 477 6 ABP72344 Brain fac
661 40 88.9 477 7 ADJ80175 Novel hum
662 40 88.9 480 2 AAR44551 Brain fac
663 40 88.9 480 4 ABG12623 Novel hum
664 40 88.9 481 5 ABB57076 Mouse isc
665 40 88.9 483 8 ADU02754 Novel hum
666 40 88.9 485 7 ADD30204 Plant yie
667 40 88.9 485 8 ADI41561 Plant tra
668 40 88.9 485 9 AEA26633 Stress to
669 40 88.9 486 8 ADN72955 Thale cre
670 40 88.9 489 8 ADQ19922 Human sof
671 40 88.9 491 8 ADY07141 Plant ful
672 40 88.9 492 8 ADE28313 Human Kpp
673 40 88.9 492 8 ADR86074 Aspergill
674 40 88.9 492 8 ADU02730 Novel hum
675 40 88.9 493 4 ABG12673 Novel hum
676 40 88.9 494 3 AAY69386 Amino aci
677 40 88.9 494 3 AAY69393 Amino aci
678 40 88.9 498 8 ADR87354 Plant ful
679 40 88.9 500 4 ABG03058 Novel hum
680 40 88.9 500 4 ABG19840 Novel hum
681 40 88.9 500 6 ABR01808 Human can

Adv68459 NHK amino
Ady54996 Human nov
Abm90394 Rice abio
Abm87638 Rice abio
Aam25490 Human pro
Abo43147 A. thalia
Ado02519 Thalecres
Aau03507 Human pro
Abb64723 Drosophil
Abu33914 Protein e
Aay17066 Human 3-O
Aab95507 Human pro
Adi09937 Human hep
Adj76166 Marker ge
Adl61251 Human hep
Adx26316 Novel cel
Aab70114 Murine 3 -
Aab70115 Human 3-O
Aab08218 A protein
Adx83401 Human TEG
Adt58544 Plant pol
Abr42686 Petunia a
Aae32027 Human kin
Adj96641 Human Nim
Aae15660 Human hom
Abp97648 Amino aci
Aeb48139 Amino aci
Aag32324 Arabidops
Aag09613 Arabidops
Adi43535 Plant tra
Aea26765 Stress to
Abp69406 Human pol
Abu65180 Human NOV
Adn62011 Human nov
Ady18827 PRO polyp
Abm91399 M. xanthu
Aag09612 Arabidops
Adf74169 Human nov
Ady25088 Plant ful
Aag49463 Arabidops
Abp75500 Human sec
Aag49462 Arabidops
Add48354 Human Pro
Abo69519 Pseudomon
Ade40443 Human cho
Aau17200 Novel sig
Adj70999 Human hea
Adb93908 Human nov
Ady08284 Plant ful
Aab41079 Human ORF
Adj71959 Human PMM
Abp72344 Brain fac
Adj80175 Novel hum
Aar44551 Brain fac
Abg12623 Novel hum
Abb57076 Mouse isc
Adu02754 Novel hum
Add30204 Plant yie
Adi41561 Plant tra
Aea26633 Stress to
Adn72955 Thale cre
Adq19922 Human sof
Ady07141 Plant ful
Ade28313 Human Kpp
Adr86074 Aspergill
Adu02730 Novel hum
Abg12673 Novel hum
Aay69386 Amino aci
Aay69393 Amino aci
Adx87354 Plant ful
Abg03058 Novel hum
Abg19840 Novel hum
Abr01808 Human can

682	40	88.9	503	7	ABR62851	Abr62851 Murine To
683	40	88.9	505	8	ADT59741	Adt59741 Plant pol
684	40	88.9	519	9	ADM17238	Adw17238 Eucalyptu
685	40	88.9	522	8	ADX91370	Adx91370 Plant ful
686	40	88.9	522	8	ADX93027	Adx93027 Plant ful
687	40	88.9	523	7	ADC87177	Adc87177 Human GPC
688	40	88.9	524	3	AAG17154	Aag17154 Arabidops
689	40	88.9	524	7	ADC31373	Adc31373 Human nov
690	40	88.9	533	8	ADX96523	Adx96523 Plant ful
691	40	88.9	535	8	ADX91150	Adx91150 Plant ful
692	40	88.9	543	6	AAE30109	Aae30109 Heliothis
693	40	88.9	546	2	AAW39788	Aaw39788 Tobacco p
694	40	88.9	547	7	ABM89005	Abm89005 Rice abio
695	40	88.9	549	8	ADX88993	Adx88993 Plant ful
696	40	88.9	553	4	ABB60817	Abb60817 Drosophil
697	40	88.9	553	9	ABM90608	Abm90608 M. xanthu
698	40	88.9	557	8	ADL57099	Adl57099 Human NOV
699	40	88.9	562	7	ADB65060	Adb65060 Human pro
700	40	88.9	562	8	ADR58973	Adr58973 Human Elk
701	40	88.9	562	8	ADR58971	Adr58971 Human Elk
702	40	88.9	570	7	ADB64564	Adb64564 Human pro
703	40	88.9	581	5	ABG66745	Abg66745 Human nov
704	40	88.9	585	4	AAM93788	Aam93788 Human pol
705	40	88.9	585	8	ADL31778	Adl31778 Human pro
706	40	88.9	585	8	ABM81730	Abm81730 Tumour-as
707	40	88.9	589	5	ABP73306	Abp73306 Candida a
708	40	88.9	594	4	ABB61362	Abb61362 Drosophil
709	40	88.9	600	8	ADO62366	Ado62366 Transcrip
710	40	88.9	603	5	AAE22088	Aae22088 Human nov
711	40	88.9	603	8	ADJ65163	Adj65163 Potassium
712	40	88.9	603	8	ADT58720	Adt58720 Plant pol
713	40	88.9	615	8	ADQ19437	Adq19437 Human sof
714	40	88.9	615	8	ABM80918	Abm80918 Tumour-as
715	40	88.9	615	9	AEA17692	Aea17692 Alzheimer
716	40	88.9	616	4	AAU09069	Aau09069 Human neu
717	40	88.9	616	6	AAO19950	Aao19950 Alzheimer
718	40	88.9	616	7	ADD14195	Add14195 Human src
719	40	88.9	616	9	ADY70363	Ady70363 Human bet
720	40	88.9	616	9	AEA17683	Aea17683 Alzheimer
721	40	88.9	617	7	ADE83413	Ade83413 Rat Prote
722	40	88.9	617	7	ADE58004	Ade58004 Rat Prote
723	40	88.9	622	4	ABB62816	Abb62816 Drosophil
724	40	88.9	622	5	ABG79656	Abg79656 Invertebr
725	40	88.9	622	8	ADJ65161	Adj65161 Potassium
726	40	88.9	622	8	AEB44574	Aeb44574 Sleep dis
727	40	88.9	639	8	ADF77438	Adf77438 MLV Env p
728	40	88.9	641	8	ADF77489	Adf77489 MLV Env p
729	40	88.9	642	4	AAE08927	Aae08927 Human NOV
730	40	88.9	642	7	ADM29403	Adm29403 Human nov
731	40	88.9	645	2	AAV28817	Aay28817 pt326_4 s
732	40	88.9	645	4	AAM78331	Aam78331 Human pro
733	40	88.9	645	4	AAM39207	Aam39207 Human pol
734	40	88.9	645	4	AAB94070	Aab94070 Human pro
735	40	88.9	650	2	AAR711382	Aar711382 Drosophil
736	40	88.9	650	8	ADP04179	Adp04179 Human col
737	40	88.9	655	3	AAV96448	Aay96448 Forkhead
738	40	88.9	655	3	AAB06076	Aab06076 Human hom
739	40	88.9	655	8	ADN06032	Adn06032 Antipsori
740	40	88.9	655	8	ADO22517	Ado22517 Biochemic
741	40	88.9	655	8	ADN40823	Adn40823 Human for
742	40	88.9	655	8	ABM81512	Abm81512 Tumour-as
743	40	88.9	655	9	ADY14882	Ady14882 PRO polyp
744	40	88.9	656	7	ADD47623	Add47623 Human Pro
745	40	88.9	656	7	ADD48792	Add48792 Human Pro
746	40	88.9	656	9	ADV99949	Adv99949 Human PEM
747	40	88.9	659	9	ADV99948	Adv99948 Human PEM
748	40	88.9	659	9	ADV99945	Adv99945 Human PEM
749	40	88.9	669	2	AAR86408	Aar86408 Human mat
750	40	88.9	669	4	AAB84617	Aab84617 Amino aci
751	40	88.9	669	4	AAE10424	Aae10424 Human mat
752	40	88.9	669	4	AAO12939	Aao12939 Human pol
753	40	88.9	690	5	AAE22154	Aae22154 Human TRN
754	40	88.9	690	6	ADA54663	Ada54663 Human pro

755	40	88.9	690	7	ADM29409	Adm29409 Human nov
756	40	88.9	690	7	ADM29397	Adm29397 Human nov
757	40	88.9	690	7	ADM29389	Adm29389 Human nov
758	40	88.9	695	7	ADM29399	Adm29399 Human nov
759	40	88.9	696	7	ADM29407	Adm29407 Human nov
760	40	88.9	706	7	ADM05234	Adm05234 Human pro
761	40	88.9	708	7	ADC19791	Adc19791 Human can
762	40	88.9	714	8	ADY06938	Ady06938 Plant ful
763	40	88.9	717	7	ADC01844	Adc01844 C. albica
764	40	88.9	721	5	AAE22152	Aae22152 Human TRN
765	40	88.9	721	7	ADM04785	Adm04785 Human pro
766	40	88.9	721	7	ADM29395	Adm29395 Human nov
767	40	88.9	737	9	ADV99974	Adv99974 Human PEM
768	40	88.9	737	9	ADV99972	Adv99972 Human PEM
769	40	88.9	738	7	ADM29393	Adm29393 Human nov
770	40	88.9	748	4	AAM40227	Aam40227 Human pol
771	40	88.9	771	4	ABB64331	Abb64331 Drosophil
772	40	88.9	775	5	AAU93161	Aau93161 Arabidops
773	40	88.9	780	6	AAE33674	Aae33674 Human str
774	40	88.9	781	7	ADM29391	Adm29391 Human nov
775	40	88.9	804	4	ABG03827	Abg03827 Novel hum
776	40	88.9	804	7	ADC33210	Adc33210 Human nov
777	40	88.9	814	8	ADK70212	Adk70212 Human oes
778	40	88.9	814	8	ADL72166	Adl72166 Human sol
779	40	88.9	817	5	AAO15419	Aao15419 Novel hum
780	40	88.9	828	4	ABB58368	Abb58368 Drosophil
781	40	88.9	834	5	ABP74067	Abp74067 Candida a
782	40	88.9	863	8	ADI43159	Adi43159 Plant tra
783	40	88.9	863	8	ADO03242	Ado03242 Thalecres
784	40	88.9	863	8	ADO62535	Ado62535 Transcrip
785	40	88.9	876	8	ABO84874	Abo84874 Human can
786	40	88.9	885	2	AAR99740	Aar99740 Human pho
787	40	88.9	914	4	ABB66394	Abb66394 Drosophil
788	40	88.9	933	9	AEA00186	Aea00186 Human TAT
789	40	88.9	933	9	AEA00706	Aea00706 Human TAT
790	40	88.9	965	7	ADC31607	Adc31607 Human nov
791	40	88.9	978	7	ADC31742	Adc31742 Human nov
792	40	88.9	1026	3	AAV83025	Aay83025 Staufen p
793	40	88.9	1026	4	ABB61775	Abb61775 Drosophil
794	40	88.9	1045	4	ABB62505	Abb62505 Drosophil
795	40	88.9	1051	2	AAV29321	Aay29321 Human ata
796	40	88.9	1052	6	AAE33773	Aae33773 Human nuc
797	40	88.9	1052	7	ADE71258	Ade71258 Novel hum
798	40	88.9	1059	8	ADN99406	Adn99406 Novel hum
799	40	88.9	1060	8	ADS10464	Adsl0464 Human the
800	40	88.9	1063	8	ADR09556	Adr09556 Human pro
801	40	88.9	1070	9	ADX05832	Adx05832 Cyclin-de
802	40	88.9	1105	8	ADT04693	Adt04693 Rice DNA
803	40	88.9	1117	8	ADI82548	Adi82548 Human mod
804	40	88.9	1134	5	AAE24341	Aae24341 Human lun
805	40	88.9	1134	7	ADF74123	Adf74123 Human nov
806	40	88.9	1232	4	ABB58856	Abb58856 Drosophil
807	40	88.9	1242	2	AAV13461	Aay13461 Amino aci
808	40	88.9	1242	2	AAW93972	Aaw93972 Human IRS
809	40	88.9	1242	4	AAB83921	Aab83921 Amino aci
810	40	88.9	1242	6	ABB99797	Abb99797 Amino aci
811	40	88.9	1242	6	ABG72369	Abg72369 Human Ins
812	40	88.9	1242	6	ABG72372	Abg72372 Human Ins
813	40	88.9	1242	6	ABG72370	Abg72370 Human Ins
814	40	88.9	1242	6	ABG72371	Abg72371 Human Ins
815	40	88.9	1242	6	ABG72368	Abg72368 Human Ins
816	40	88.9	1242	6	ADA00642	Ada00642 Human IRS
817	40	88.9	1242	8	ADJ66659	Adj66659 IRS1 prot
818	40	88.9	1242	8	ADN98329	Adn98329 Human ins
819	40	88.9	1242	8	ADO44347	Ado44347 Human IRS
820	40	88.9	1242	9	ADU87427	Adu87427 Human IRS
821	40	88.9	1242	9	ADX07852	Adx07852 Cyclin-de
822	40	88.9	1242	9	ADX86115	Adx86115 Human ins
823	40	88.9	1242	9	ADZ79340	Adz79340 Human ins
824	40	88.9	1242	9	AEA81147	Aea81147 Human ins
825	40	88.9	1243	2	AAR67708	Aar67708 Insulin r
826	40	88.9	1243	7	ABM85838	Abm85838 Mouse pro
827	40	88.9	1297	7	ABM90342	Abm90342 Rice abio

828	40	88.9	1316	4	ABG222997	Abg222997	Novel hum
829	40	88.9	1316	7	ADK40903	Adk40903	Novel hum
830	40	88.9	1316	8	ADR15627	Adrl5627	Kinase 36
831	40	88.9	1316	9	AEB32334	Aeb32334	Human pro
832	40	88.9	1316	9	AEB32358	Aeb32358	Human pro
833	40	88.9	1321	4	ABG25416	Abg25416	Novel hum
834	40	88.9	1437	7	ADJ69580	Adj69580	Human hea
835	40	88.9	1437	7	ADJ69918	Adj69918	Human hea
836	40	88.9	1439	7	ABM88861	Abm88861	Rice abio
837	40	88.9	1640	7	ADC59314	Adc59314	Human pol
838	40	88.9	1703	4	ABB66223	Abb66223	Drosophil
839	40	88.9	1798	3	AAy51611	Aay51611	Human HSG
840	40	88.9	1819	7	ABM85839	Abm85839	Human pro
841	40	88.9	1963	4	ABB62819	Abb62819	Drosophil
842	40	88.9	1991	4	ABB60651	Abb60651	Drosophil
843	40	88.9	2075	4	ABB61677	Abb61677	Drosophil
844	40	88.9	2464	4	ABB63174	Abb63174	Drosophil
845	40	88.9	2566	8	ADS10740	Ads10740	Human the
846	40	88.9	3119	2	AAW36888	Aaw36888	Mouse Hun
847	40	88.9	3119	2	AAW44743	Aaw44743	Mouse hun
848	40	88.9	3139	2	AAy08898	Aay08898	Human Hun
849	40	88.9	3238	4	ABB71715	Abb71715	Drosophil
850	40	88.9	4019	4	AAE13839	Aae13839	Human lun
851	40	88.9	4019	7	ADD66733	Add66733	Human lun
852	40	88.9	4019	7	ADE87987	Ade87987	Human lun
853	40	88.9	4025	5	ABP69736	Abp69736	Human pol
854	40	88.9	4025	8	ADU18074	Adul18074	Human can
855	40	88.9	19938	6	ABP76681	Abp76681	Streptomy
856	39	86.7	9	8	ADU03983	Adu03983	HTLV-I Ga
857	39	86.7	10	8	ADU03985	Adu03985	HTLV-I Ga
858	39	86.7	12	8	ADU03991	Adu03991	HTLV-I Ga
859	39	86.7	13	8	ADU03987	Adu03987	HTLV-I Ga
860	39	86.7	15	2	AAW38957	Aaw38957	Peptide r
861	39	86.7	15	8	ADU03986	Adu03986	HTLV-I Ga
862	39	86.7	15	8	ADT90816	Adt90816	PAG2 prot
863	39	86.7	16	2	AAW38947	Aaw38947	Peptide r
864	39	86.7	28	2	AAR39867	Aar39867	E2 peptid
865	39	86.7	34	8	ABO54390	Abo54390	Human gen
866	39	86.7	43	8	ADU04374	Adu04374	HTLV-I Ga
867	39	86.7	50	2	AAy29804	Aay29804	Human GAB
868	39	86.7	51	6	ADA48234	Ada48234	Rice prot
869	39	86.7	59	3	AAG02396	Aag02396	Human sec
870	39	86.7	61	3	AAG09627	Aag09627	Arabidops
871	39	86.7	62	8	ABO58404	Abo58404	Human gen
872	39	86.7	67	4	AAm16772	Aam16772	Peptide #
873	39	86.7	67	4	ABB35758	Abb35758	Peptide #
874	39	86.7	67	4	AAM29259	Aam29259	Peptide #
875	39	86.7	67	4	ABB30592	Abb30592	Peptide #
876	39	86.7	67	4	ABB21182	Abb21182	Protein #
877	39	86.7	67	4	AAm68948	Aam68948	Human bon
878	39	86.7	67	4	AAm56566	Aam56566	Human bra
879	39	86.7	67	4	ABG50611	Abg50611	Human liv
880	39	86.7	67	4	AAm04489	Aam04489	Peptide #
881	39	86.7	67	5	ABG38532	Abg38532	Human pep
882	39	86.7	68	3	AAy84596	Aay84596	Fragment
883	39	86.7	68	3	AAG53874	Aag53874	Arabidops
884	39	86.7	69	7	ADC32722	Adc32722	Human nov
885	39	86.7	70	3	AAG10803	Aag10803	Arabidops
886	39	86.7	72	2	AAR41429	Aar41429	HTLV-I ga
887	39	86.7	74	3	AAG53873	Aag53873	Arabidops
888	39	86.7	75	4	AAO50005	Aao50005	Human pol
889	39	86.7	76	3	AAG10802	Aag10802	Arabidops
890	39	86.7	77	9	ADV76639	Adv76639	Human TER
891	39	86.7	78	3	AAG53872	Aag53872	Arabidops
892	39	86.7	83	4	AAm95449	Aam95449	Human rep
893	39	86.7	83	5	ABP09336	Abp09336	Human ORF
894	39	86.7	90	8	ADX78974	Adx78974	Plant ful
895	39	86.7	95	3	AAG58354	Aag58354	Arabidops
896	39	86.7	99	8	ADT57219	Adt57219	Plant pol
897	39	86.7	104	4	AAm13657	Aam13657	Peptide #
898	39	86.7	104	4	ABB32587	Abb32587	Peptide #
899	39	86.7	104	4	AAm26054	Aam26054	Peptide #
900	39	86.7	104	4	ABB27438	Abb27438	Human pep

901	39	86.7	104	4	ABB18083	Abb18083	Protein #
902	39	86.7	104	4	AAm65795	Aam65795	Human bon
903	39	86.7	104	4	AAm53415	Aam53415	Human bra
904	39	86.7	104	4	ABG47440	Abg47440	Human liv
905	39	86.7	104	4	AAm01407	Aam01407	Peptide #
906	39	86.7	104	5	ABG35427	Abg35427	Human pep
907	39	86.7	107	3	AAy84598	Aay84598	Fragment
908	39	86.7	107	4	AAO12356	Aao12356	Human pol
909	39	86.7	109	4	AAO02138	Aao02138	Human pol
910	39	86.7	110	4	AAO01457	Aao01457	Human pol
911	39	86.7	110	8	ABO56225	Abo56225	Human gen
912	39	86.7	119	4	AAO07689	Aao07689	Human pol
913	39	86.7	126	3	AAG56294	Aag56294	Arabidops
914	39	86.7	126	8	ADT56036	Adt56036	Plant pol
915	39	86.7	127	5	AAE29154	Aae29154	Human non
916	39	86.7	128	8	ADT59501	Adt59501	Plant pol
917	39	86.7	131	9	ABM91613	Abm91613	M. xanthu
918	39	86.7	135	5	ADK34119	Adk34119	Novel hum
919	39	86.7	136	4	AAO00750	Aao00750	Human pol
920	39	86.7	139	3	AAy45011	Aay45011	Partial h
921	39	86.7	140	3	AAy84588	Aay84588	A third p
922	39	86.7	140	3	AAy68707	Aay68707	Amino aci
923	39	86.7	140	3	AAy68711	Aay68711	Amino aci
924	39	86.7	140	5	ABB82396	Abb82396	Human neu
925	39	86.7	140	5	AAO22937	Aao22937	Human pre
926	39	86.7	140	5	ABJ15110	Abj15110	Pre-pro-
927	39	86.7	140	8	ADR16440	Adr16440	Neublasti
928	39	86.7	140	8	ADU46759	Adu46759	Human neu
929	39	86.7	140	9	ADV35156	Adv35156	Human neu
930	39	86.7	142	8	ADX72424	Adx72424	Plant ful
931	39	86.7	147	8	ADY23420	Ady23420	Plant ful
932	39	86.7	150	5	ABP64467	Abp64467	Human ORF
933	39	86.7	153	7	ABM86340	Abm86340	Rice abio
934	39	86.7	158	8	ADY23419	Ady23419	Plant ful
935	39	86.7	158	9	ADV35173	Adv35173	Signal pe
936	39	86.7	158	9	ABM92717	Abm92717	M. xanthu
937	39	86.7	159	3	AAy44774	Aay44774	Protein-2
938	39	86.7	159	9	ADV35166	Adv35166	Immunoglo
939	39	86.7	159	9	ADV35180	Adv35180	Signal pe
940	39	86.7	161	3	AAG58353	Aag58353	Arabidops
941	39	86.7	164	8	ABO60055	Abo60055	Human gen
942	39	86.7	165	2	AAy08459	Aay08459	Rat Ptx1
943	39	86.7	165	8	ADX79099	Adx79099	Plant ful
944	39	86.7	166	9	ADV35187	Adv35187	Signal pe
945	39	86.7	168	4	ABG13945	Abg13945	Novel hum
946	39	86.7	171	3	AAy71561	Aay71561	Rat Nogo
947	39	86.7	178	4	AAm25955	Aam25955	Human pro
948	39	86.7	178	7	ADC33057	Adc33057	Human nov
949	39	86.7	178	8	ADM87723	Adm87723	Human EST
950	39	86.7	179	9	ADV35160	Adv35160	Signal pe
951	39	86.7	181	9	ADM45341	Adm45341	Human art
952	39	86.7	181	9	ADM45344	Adm45344	Human art
953	39	86.7	181	9	ADM45340	Adm45340	Human art
954	39	86.7	181	9	ADM45345	Adm45345	Human art
955	39	86.7	181	9	ADM45343	Adm45343	Human art
956	39	86.7	181	9	ADM45342	Adm45342	Human art
957	39	86.7	185	8	ADQ16418	Adq16418	Nucleotid
958	39	86.7	188	4	AAG82812	Aag82812	S. epider
959	39	86.7	188	7	ABM86285	Abm86285	Rice abio
960	39	86.7	188	9	ABM90756	Abm90756	M. xanthu
961	39	86.7	198	5	ABP38468	Abp38468	Staphyloc
962	39	86.7	198	8	ADS04629	Ads04629	Staphyloc
963	39	86.7	202	7	AAO24273	Aao24273	Human vas
964	39	86.7	205	6	ABU20105	Abu20105	Protein e
965	39	86.7	209	2	AAW26615	Aaw26615	Ferredoxi
966	39	86.7	211	7	ADT88972	Adt88972	Tobacco M
967	39	86.7	213	5	ABB90623	Abb90623	Chlamydia
968	39	86.7	213	8	ADX94802	Adx94802	Plant ful
969	39	86.7	213	9	ADM65020	Adm65020	C. pneumo
970	39	86.7	215	8	ADY11803	Ady11803	Plant ful
971	39	86.7	216	2	AAy35299	Aay35299	Chlamydia
972	39	86.7	220	3	AAy84583	Aay84583	Amino aci
973	39	86.7	220	3	AAy44776	Aay44776	Short spl

974 39 86.7 220 3 AAY68710 Aay68710 A human p
975 39 86.7 220 4 AAB50978 Aab50978 Human PRO
976 39 86.7 220 5 AAU86158 Aau86158 Human PRO
977 39 86.7 220 5 ABB84975 Abb84975 Human PRO
978 39 86.7 220 5 ABG30698 Abg30698 Human art
979 39 86.7 220 5 ABB82388 Abb82388 Human neu
980 39 86.7 220 5 ABB95581 Abb95581 Human ang
981 39 86.7 220 5 AAO22940 Aao22940 Human foe
982 39 86.7 220 6 ABU56702 Abu56702 Lung canc
983 39 86.7 220 6 ABU56539 Abu56539 Lung canc
984 39 86.7 220 6 ABU56703 Abu56703 Lung canc
985 39 86.7 220 6 ABU56540 Abu56540 Lung canc
986 39 86.7 220 6 ABU71444 Abu71444 Human neo
987 39 86.7 220 7 ADD10607 Add10607 Human sec
988 39 86.7 220 7 ADD11567 Add11567 Human sec
989 39 86.7 220 7 ADD37360 Add37360 Human sec
990 39 86.7 220 7 ADJ37343 Adj37343 Human tum
991 39 86.7 220 7 ADN39086 Adn39086 Cancer/an
992 39 86.7 220 7 ADN39084 Adn39084 Cancer/an
993 39 86.7 220 8 ADE41568 Ade41568 Human sec
994 39 86.7 220 8 ADG68267 Adg68267 Human PRO
995 39 86.7 220 8 ADH43751 Adh43751 Human PRO
996 39 86.7 220 8 ADK83096 Adk83096 Human PRO
997 39 86.7 220 8 ADR16439 Adr16439 Human pre
998 39 86.7 220 8 ADU46758 Adu46758 Human pre
999 39 86.7 220 9 ADV35145 Adv35145 Human pre
1000 39 86.7 220 9 ADZ80734 Adz80734 Amino aci

ALIGNMENTS

RESULT 1
AAB17249
ID AAB17249 standard; peptide; 7 AA.
XX
AC AAB17249;
XX
DT 31-OCT-2000 (first entry)
XX
DE SH3 antagonist peptide sequence SEQ ID NO:305.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CRUA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 39; Page 302; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an

CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 45; DB 3; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
Db 1 RPQPPPP 7

RESULT 2
ABB73242
ID ABB73242 standard; peptide; 7 AA.
XX
AC ABB73242;
XX
DT 05-APR-2002 (first entry)
XX
DE Src homology3 (SH3) antagonist peptide SEQ ID NO:305.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW anastnaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX

PS Claim 39; Page 55; 176pp; English.

XX The present invention describes a vehicle-peptide molecule (I) or its

CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,

CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,

CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and

CC neuroprotective activities. (I) can be used as a therapeutic or

CC prophylactic agent as well as for screening purposes. (I) is useful for

CC diagnosing diseases characterised by dysfunction of their associated

CC protein of interest, for identifying normal or abnormal proteins of

CC interest, as a part of diagnostic kit to detect the presence of their

CC proteins of interest in a biological sample. Additionally, (I) is useful

CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,

CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,

CC infertility, and neurological degenerative diseases. (I), comprising EPO-

CC mimetic compounds are useful for treating disorders characterised by low

CC red blood cell levels such as anaemia. The TPO-mimetic comprising

CC compounds are useful for treating conditions that involve an existing

CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet

CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic

CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,

CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777

CC represent amino acid and nucleic acid sequences used in the

CC exemplification of the present invention

XX

SQ Sequence 7 AA;

Query Match 100.0%; Score 45; DB 5; Length 7;

Best Local Similarity 100.0%; Pred. No. 2e+06;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7

Db |||||

1 RPQPPPP 7

RESULT 3

ADJ73396

ID ADJ73396 standard; peptide; 7 AA.

XX

AC ADJ73396;

XX

DT 06-MAY-2004 (first entry)

XX

DE SH3 antagonist peptide sequence SeqID 851.

XX

KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;

KW cardiovascular; infectious; malignant; neurologic disease; anaemia;

KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;

KW SH3.

XX

OS Synthetic.

XX

PN WO2003084477-A2.

XX

PD 16-OCT-2003.

XX

PF 24-MAR-2003; 2003WO-US009139.

XX

PR 29-MAR-2002; 2002US-0368791P.

XX

PA (CENZ) CENTOCOR INC.

XX

PI Heavner GA, Knight DM, Scallion BJ, Ghrayeb J;

XX

DR WPI; 2003-804237/75.

XX

XX New CDR mimetibody comprising a portion of a heavy or light chain

PT variable region comprising human framework or ligand binding region,

PT useful for preparing a composition for treating e.g., immune,

PT cardiovascular or neurologic disease.

XX

PS Disclosure; SEQ ID NO 851; 97pp; English.

XX This invention relates to novel mammalian CDR mimetibodies, specific

CC portions or variants thereof. Specifically, it refers to an antibody

CC fragment where a protein has been inserted into, or replaces a portion

CC of, one or more CDR regions, such that each CDR mimetibody comprises at

CC least one portion of a heavy chain or light chain variable region, which

CC itself comprises at least one human framework region and at least one

CC ligand binding region (LBR). The present invention describes human

CC mimetibodies, including modified immunoglobulins and cleavage products

CC that can be useful in gene therapy and the generation of transgenic

CC plants and animals. Furthermore, the CDR mimetibody is useful for

CC preparing compositions for modulating, treating or reducing the symptoms

CC of immune, cardiovascular, infectious, malignant and/ or neurologic

CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,

CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This

CC peptide sequence is an SH3 antagonist peptide sequence used to make a

CC mimetibody of the invention.

XX

SQ Sequence 7 AA;

Query Match 100.0%; Score 45; DB 7; Length 7;

Best Local Similarity 100.0%; Pred. No. 2e+06;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7

Db |||||

1 RPQPPPP 7

RESULT 4

ADJ53030

ID ADJ53030 standard; peptide; 7 AA.

XX

AC ADJ53030;

XX

DT 06-MAY-2004 (first entry)

XX

DE CH1 deleted mimetibody-related peptide SeqID851.

XX

KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;

KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;

KW fungicide; gene therapy; immune disorder; cardiovascular disease;

KW arrhythmia; hypertension; heart failure; neurodegenerative;

KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;

KW cancerous condition; infectious disease; bacterial infection;

KW viral infection; fungal infection.

XX

OS Unidentified.

OS Synthetic.

XX

PN WO2004002417-A2.

XX

PD 08-JAN-2004.

XX

PF 27-JUN-2003; 2003WO-US020347.

XX

PR 28-JUN-2002; 2002US-0392431P.

XX

PA (CENZ) CENTOCOR INC.

XX

PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;

PI Kutoloski KA;

XX

DR WPI; 2004-082870/08.

XX

XX New CH1-deleted mimetibody polypeptides and nucleic acids, useful for

PT modulating, treating, alleviating, preventing an immune, cardiovascular,

PT or neurodegenerative disease or disorder, anemia, cancer, or infectious

PT diseases.

XX

PS Claim 3; SEQ ID NO 851; 129pp; English.

XX

CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences

CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.

XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 45; DB 8; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
Db 1 RPQPPPP 7

RESULT 5
ADJ51991
ID ADJ51991 standard; peptide; 7 AA.

XX AC ADJ51991;

XX DT 06-MAY-2004 (first entry)

XX DE CH1 deleted mimetibody-related peptide SeqID851.

XX KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmologic; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunologic disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.

XX OS Unidentified.
OS Synthetic.

XX PN WO2004002424-A2.

XX PD 08-JAN-2004.

XX PF 30-JUN-2003; 2003WO-US020495.

XX PR 28-JUN-2002; 2002US-0392431P.

XX PR 19-SEP-2002; 2002US-0412144P.

XX PA (CENZ) CENTOCOR INC.

XX PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;

XX DR WPI; 2004-082872/08.

XX PT New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and

PT nutritional disorders.

XX Claim 15; SEQ ID NO 851; 123pp; English.

XX CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC antiinflammatory, neuroleptic, ophthalmologic, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CH1 deleted mimetibody, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunologic, allergic, infectious,
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetibody of the invention.

XX SQ Sequence 7 AA;

Query Match 100.0%; Score 45; DB 8; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7

Db 1 RPQPPPP 7

RESULT 6

AAW11113

ID AAW11113 standard; peptide; 13 AA.

XX AC AAW11113;

XX DT 25-JUN-1997 (first entry)

XX DE Src SH3 domain-binding peptide used in signal transduction modulation.

XX KW Src; SH3; Src homology region 3; binding affinity; oncogenic protein;
KW protein tyrosine kinase; signal transduction; RNA processing;
KW trafficking; translation.

XX OS Synthetic.

XX PN WO9603649-A1.

XX PD 08-FEB-1996.

XX PF 24-JUL-1995; 95WO-US009382.

XX PR 22-JUL-1994; 94US-00278865.

XX PR 07-JUN-1995; 95US-00483555.

XX PA (UYNC-) UNIV NORTH CAROLINA.

XX PI Sparks AB, Kay BK, Thorn JM, Quilliam LA, Der CJ;

XX DR WPI; 1996-117151/12.

XX PT Peptide with binding affinity for Src homology region 3 (SH3) domains of
PT proteins - useful for e.g. modulating signal transduction pathways at the
PT cellular level, esp. protein tyrosine kinase-mediated.

XX PS Claim 38; Page 87; 116pp; English.

XX CC AAW11098-W11124 are peptides that bind to the Src SH3 domain. The SH3

CC binding peptides are useful in modulating signal transduction pathways at
CC the cellular level (especially protein tyrosine kinase-mediated),
CC modulating oncogenic protein activity, or providing compounds for the
CC development of drugs with the ability to modulate broad classes, as well
CC as specific classes, of proteins involved in signal transduction and also
CC for regulating the processing, trafficking or translation of RNA.
CC Conjugates of the peptides with detectable labels or imaging agents are
CC useful for imaging cells, tissues and organs in which Src or Src-related
CC proteins are expressed
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 45; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 4 RPQPPPP 10

RESULT 7
AAW25511
ID AAW25511 standard; peptide; 31 AA.
XX
AC AAW25511;

XX 27-MAR-1998 (first entry)

DE Random peptide recombinant clone R8C.YES3.9.

XX Cortactin; SH3 domain; binding peptide; Src homology region 3;
KW tyrosine kinase; immune response; lymphokine; interleukin 1; Nck; Abl;
KW PLCgamma; p53bp2; Crk; Yes; Grb2.

XX Synthetic.
OS Unidentified.

XX WO9730074-A1.

XX 21-AUG-1997.

PF 14-FEB-1997; 97WO-US002298.

XX 16-FEB-1996; 96US-00602999.

XX (CYTO-) CYTOGEN CORP.
PA (UYNC-) UNIV NORTH CAROLINA.

XX Sparks AB, Kay BK, Thorn JM, Quilliam LA, Der CJ, Fowlkes DM;
PI Rider JE;

XX WPI; 1997-424972/39.

XX Src homology region 3 binding peptide - used to activate Src tyrosine
PT kinase(s) and to stimulate immune response by increasing production of
PT certain lymphokine(s), e.g. interleukin-1.

XX Disclosure; Fig 5; 131pp; English.

XX The present sequence represents a random peptide recombinant isolated by
CC the method of the present invention. SH3 (Src homology region 3) binding
CC peptides are selected from: (a) peptides which bind the SH3 domain of
CC Cortactin; (b) peptides which bind the middle SH3 domain of Nck; (c)
CC peptides which bind the SH3 domain of Abl; (d) peptides which bind the
CC SH3 domain of Src; (e) peptides which bind the SH3 domain of PLC gamma;
CC (f) peptides which bind the SH3 domain of p53bp2; (g) peptides which bind
CC the amino-terminal SH3 domain of Crk; (h) peptides which bind the SH3
CC domain of Yes; and (i) peptides which bind the amino-terminal SH3 domain
CC of Grb2. The purified binding peptides can be used in the method to
CC identify inhibitors of their binding to their respective SH3 domains,
CC which could be used to modulate the pharmacological activity of proteins
CC or polypeptide containing the SH3 domain. The peptides can also be used

CC to activate Src or Src-related protein tyrosine kinases, to stimulate the
CC immune response by increasing the production of certain lymphokines, e.g.
CC tumour necrosis factor-alpha and interleukin-1, or to deliver a
CC conjugated molecule to certain cellular compartments containing Src or
CC Src related proteins

SQ Sequence 31 AA;

Query Match 100.0%; Score 45; DB 2; Length 31;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 18 RPQPPPP 24

RESULT 8
ADI45407
ID ADI45407 standard; protein; 122 AA.
XX
AC ADI45407;

XX 22-APR-2004 (first entry)

DE Rice isoprenoid biosynthesis-associated protein #169.

XX Rice; isoprenoid biosynthesis; plant; isopentenyl diphosphate; IPP;
KW dimethylallyl alcohol; DMAPP; short-chain plastid prenyltransferase;
KW gibberellin; carotenoid; abscisic acid; tocopherol; plastoquinone;
KW phylloquinone; mevalonate pathway; phytosterol; brassinosteroid;
KW ubiquinone; monoterpene; sesquiterpene; protein prenylation; chlorophyll;
KW haeme; yield.

XX Oryza sativa.

XX US2004010815-A1.

XX 15-JAN-2004.

PF 26-SEP-2002; 2002US-00259194.

XX 26-SEP-2001; 2001US-0325277P.

PR 04-APR-2002; 2002US-0370620P.

PR 04-APR-2002; 2002US-0370743P.

XX (LANG/) LANGE B M.

PA (GHAS/) GHASSEMIAN M.

PA (BRIG/) BRIGGS S P.

PA (COOP/) COOPER B.

PA (GLAZ/) GLAZEBROOK J.

PA (GOFF/) GOFF S A.

PA (KATA/) KATAGIRI F.

PA (KREP/) KREPS J.

PA (MOUG/) MOUGHAMER T.

PA (PROV/) PROVART N.

PA (RICK/) RICHE D.

PA (ZHUT/) ZHU T.

XX Lange BM, Ghassemian M, Briggs SP, Cooper B, Glazebrook J;

PI Goff SA, Katagiri F, Kreps J, Moughamer T, Provart N, Ricke D;

PI Zhu T;

XX WPI; 2004-090562/09.

DR N-PSDB; ADI45406.

XX New isolated polynucleotides and polypeptides associated with isoprenoid
PT synthesis in plants, useful for producing transgenic plants, for targeted
PT gene disruption, as well as markers or probes.
XX Claim 4; SEQ ID NO 338; 117pp; English.
XX The invention relates to a polynucleotide (or its complement, protein
CC

CC encoding fragment or reverse complement), comprising a nucleotide
CC sequence encoding a polypeptide comprising an amino acid sequence
CC involved in or associated with the biosynthesis of isoprenoids in a rice
CC plant. Also included are an isolated polypeptide involved in or
CC associated with the biosynthesis of isoprenoids in a plant, an expression
CC cassette comprising the polynucleotide, a host cell comprising the
CC expression cassette, and a transgenic plant comprising the expression
CC cassette. The polypeptides and polynucleotides include those associated
CC with the biosynthesis of isopentenyl diphosphate (IPP) and dimethylallyl
CC alcohol (DMAPP), the biosynthesis of short-chain plastid
CC prenyltransferases, the biosynthesis of gibberellins, the biosynthesis of
CC carotenoids and/or abscisic acids, the biosynthesis of tocopherols,
CC plastoquinone and/or phyllloquinone biosynthesis, the mevalonate pathway,
CC phytosterol and brassinosteroid metabolism, biosynthesis of ubiquinone,
CC biosynthesis of monoterpenes and sesquiterpenes, protein prenylation, and
CC biosynthesis of chlorophyll or haeme. Also disclosed are banana, wheat
CC and corn homologues of some of the rice polynucleotides. The
CC polynucleotides are useful for producing transgenic plants, where the
CC genome is augmented by a nucleic acid molecule of the invention, or in
CC which the corresponding gene has been disrupted, e.g. to result in a
CC loss, a decrease or an alteration in the function of the product encoded
CC by the gene. The plants may also have increased yields and/or produce a
CC better quality product than the corresponding wild-type plant. The
CC nucleic acid molecules are useful for targeted gene disruption, as well
CC as markers and probes. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20040010815. The present sequence
CC represents a Rice isoprenoid biosynthesis-associated protein of the
CC invention.

XX Sequence 122 AA;

Query Match 100.0%; Score 45; DB 8; Length 122;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
Db 113 RPQPPPP 119

RESULT 9
ABB97921
ID ABB97921 standard; protein; 173 AA.
XX
AC ABB97921;
XX
DT 03-OCT-2002 (first entry)
XX
DE Human secretory polypeptide (SPTM) 173.

XX Human; secretory protein; secretory polynucleotides; SPTM;
KW SPTM-related disease; somatic gene therapy; germline gene therapy;
KW severe combined immunodeficiency; intracellular parasite protection;
KW fungal parasite; protozoan parasite; cell proliferative disorder; cancer;
KW immune disorder; AIDS; neurological disorder; Parkinson's disease;
KW motor neuron disorder; demyelinating disease; multiple sclerosis;
KW meningitis; abscess; prion diseases; cerebral palsy;
KW neuroskeletal disorder; peripheral nervous system disorder;
KW dermatomyositis; polymyositis; myopathy; myasthenia gravis;
KW mental disorder; Tourette's syndrome.

OS Homo sapiens.
XX
XX WO200220756-A2.
PN
XX
PD 14-MAR-2002.
XX
XX 30-AUG-2001; 2001WO-US027297.
PF
XX
PR 05-SEP-2000; 2000US-0229747P.
PR 05-SEP-2000; 2000US-0229748P.

PR 05-SEP-2000; 2000US-0229749P.
PR 05-SEP-2000; 2000US-0229750P.
PR 05-SEP-2000; 2000US-0229751P.
PR 05-SEP-2000; 2000US-0230016P.
PR 05-SEP-2000; 2000US-0230583P.
PR 06-SEP-2000; 2000US-0230505P.
PR 06-SEP-2000; 2000US-0230514P.
PR 06-SEP-2000; 2000US-0230515P.
PR 06-SEP-2000; 2000US-0230517P.
PR 06-SEP-2000; 2000US-0230518P.
PR 06-SEP-2000; 2000US-0230519P.
PR 06-SEP-2000; 2000US-0230595P.
PR 06-SEP-2000; 2000US-0230596P.
PR 06-SEP-2000; 2000US-0230597P.
PR 06-SEP-2000; 2000US-0230599P.
PR 06-SEP-2000; 2000US-0230610P.
PR 06-SEP-2000; 2000US-0230864P.
PR 06-SEP-2000; 2000US-0230865P.
PR 06-SEP-2000; 2000US-0230988P.
PR 06-SEP-2000; 2000US-0230989P.
PR 06-SEP-2000; 2000US-0230990P.
PR 07-SEP-2000; 2000US-0230896P.
PR 07-SEP-2000; 2000US-0230897P.
PR 07-SEP-2000; 2000US-0230951P.
PR 07-SEP-2000; 2000US-0231163P.
PR 07-SEP-2000; 2000US-02311832P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX

PI Stuart J, Lincoln SE, Altus CM, Dufour GE, Chalup MS, Hillman JL;
PI Jones AL, Yu JY, Wright RJ, Gietzen D, Liu TF, Yap PE, Dahl CR;
PI Momiyama MG, Bradley DL, Rohatgi SD, Harris B, Roseberry AM;
PI Gerstin EH, Peralta CH, David MH, Panzer SR, Flores V, Daffo A;
PI Marwaha R, Chen AJ, Chang SC, Au AP, Inman RR;
XX

DR WPI; 2002-315658/35.
N-PSDB; ABL99918.

PT Polynucleotide sequences encoding human secretory proteins useful for
PT gene therapy of e.g. genetic deficiency disorders, cancers, and diseases
PT caused by intracellular parasites.

PS Claim 29; Page 479; 585pp; English.

XX The invention comprises the amino acid and coding sequences of human
CC secretory (SPTM) proteins. The SPTM DNA and amino acid sequences are
CC useful for treating a disease or condition associated with the expression
CC of functional SPTM. The SPTM DNA sequences are useful for somatic or
CC germline gene therapy to correct a genetic deficiency (e.g. severe
CC combined immunodeficiency). The SPTM DNA sequences are also useful in
CC providing protection against intracellular parasites (e.g. fungal
CC parasites and protozoan parasites). The SPTM DNA and protein sequences
CC are also useful for diagnosing cell proliferative disorders, cancer,
CC immune disorders (e.g. AIDS), neurological disorders (e.g. Parkinson's
CC disease), motor neuron disorders, demyelinating diseases (e.g. multiple
CC sclerosis), meningitis, abscesses, prion diseases, cerebral palsy,
CC neuroskeletal disorders, peripheral nervous system disorders,
CC dermatomyositis and polymyositis, myopathy, myasthenia gravis, and mental
CC disorders (e.g. Tourette's syndrome). Amino acid sequences ABB97749 -
CC ABB97933 represent human secretory proteins of the invention

XX SQ Sequence 173 AA;

Query Match 100.0%; Score 45; DB 5; Length 173;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
Db 106 RPQPPPP 112

RESULT 10

ABO72442
ID ABO72442 standard; protein; 249 AA.
XX
AC ABO72442;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polypeptide #4617.
XX
KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
PN US6551795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
DR WPI; 2003-615309/58.
DR N-PSDB; ABD06013.
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 21188; 455pp; English.
XX
CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 249 AA;

Query Match 100.0%; Score 45; DB 7; Length 249;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
Db 43 RPQPPPP 49
|||||||

RESULT 11
ADP29798
ID ADP29798 standard; protein; 276 AA.
XX
AC ADP29798;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #565.
XX

KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.


```
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493373P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
PA
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
XX Halenbeck RF, Huang MM, Kothakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
PI
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
PT
XX Claim 1; SEQ ID NO 1796; 428pp; English.
PS
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.
XX
XX Sequence 276 AA;
SQ
Query Match 100.0%; Score 45; DB 8; Length 276;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 RPQPPPP 7
Db 188 RPQPPPP 194
RESULT 12
AAG13899
ID AAG13899 standard; protein; 288 AA.
XX
AC AAG13899;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 13560.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493373P.
PR 08-AUG-2003; 2003US-0493577P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
```

PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-01443352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.

PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 100.0%; Score 45; DB 3; Length 288;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
|||
Db 53 RPQPPPP 59

RESULT 13
ABM97421

ID ABM97421 standard; protein; 314 AA.

XX AC ABM97421;

XX DT 02-JUN-2005 (first entry)

XX DE M. xanthus protein sequence, seq id 16620.

XX KW Transgenic plant; DNA replication; gene regulation; gene expression.

XX OS Myxococcus xanthus.

XX PN US6833447-B1.

XX PD 21-DEC-2004.

XX PF 10-JUL-2001; 2001US-00902540.

XX PR 10-JUL-2000; 2000US-0217883P.

XX PA (MONS) MONSANTO TECHNOLOGY LLC.

XX PI Goldman BS, Hinkle GJ, Slater SC, Wiegand RC;

XX DR WPI; 2005-028716/03.

XX PT New substantially purified Myxococcus xanthus nucleic acid molecule
PT encoding a nitrite reductase, useful for determining gene expression,
PT identifying mutations in a gene of interest, and for constructing
PT mutations in a gene of interest.

XX PS Example 2; SEQ ID NO 16620; 25pp; English.

XX CC The invention relates to a substantially purified nucleic acid molecule
CC encoding a nitrite reductase of SEQ ID NO. 11926. Further disclosed is a
CC recombinant DNA construct for expression of a nitrite reductase gene in a
CC plant cell, and a plant cell comprising the recombinant DNA construct.

CC The nucleic acid is useful for determining gene expression, identifying
CC mutations in a gene of interest, and for constructing mutations in a gene
CC of interest. Sequences given in records for SEQ IDs 9692-16825 represent
CC a group of 7134 Myxococcus xanthus proteins and peptides. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
XX
SQ Sequence 314 AA;

Query Match 100.0%; Score 45; DB 9; Length 314;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
| | | | |
Db 8 RPQPPPP 14

RESULT 14
AAG13898
ID AAG13898 standard; protein; 322 AA.

XX
AC AAG13898;

XX
DT 17-OCT-2000 (first entry)

XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 13559.

XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

OS Arabidopsis thaliana.

XX
PN EP1033405-A2.

XX
PD 06-SEP-2000.

XX
PF 25-FEB-2000; 2000EP-00301439.

XX
25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 23-APR-1999; 99US-0130891P.

PR 28-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132407P.

PR 04-MAY-1999; 99US-0132484P.

PR 05-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.

PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.

PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.
PR	04-AUG-1999;	99US-0147302P.		Query Match	100.0%; Score 45; DB 3; Length 322;
PR	05-AUG-1999;	99US-0147192P.		Best Local Similarity	100.0%; Pred. No. 2.4e+02;
PR	05-AUG-1999;	99US-0147260P.		Matches	7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			
PR	09-AUG-1999;	99US-0147493P.	Qy	1 RPQPPPP	7
PR	09-AUG-1999;	99US-0147935P.			
PR	10-AUG-1999;	99US-0148171P.	Db	87 RPQPPPP	93
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			
PR	13-AUG-1999;	99US-0148684P.			
PR	16-AUG-1999;	99US-0149368P.	RESULT 15		
PR	17-AUG-1999;	99US-0149175P.	AAG06371		
PR	18-AUG-1999;	99US-0149426P.	ID	AAG06371	standard; protein; 340 AA.
PR	20-AUG-1999;	99US-0149722P.	XX		
PR	20-AUG-1999;	99US-0149723P.	AC	AAG06371;	
PR	20-AUG-1999;	99US-0149929P.	XX		
PR	23-AUG-1999;	99US-0149902P.	DT	17-OCT-2000	(first entry)
PR	23-AUG-1999;	99US-0149930P.	XX		
PR	25-AUG-1999;	99US-0150566P.	DE	Arabidopsis thaliana	protein fragment SEQ ID NO: 3120.
PR	26-AUG-1999;	99US-0150884P.	XX		
PR	27-AUG-1999;	99US-0151065P.	KW	Protein identification; signal transduction pathway; metabolic pathway;	
PR	27-AUG-1999;	99US-0151066P.	KW	hybridisation assay; genetic mapping; gene expression control; promoter;	
PR	27-AUG-1999;	99US-0151066P.	XX	termination sequence.	
PR	27-AUG-1999;	99US-0151080P.	OS	Arabidopsis thaliana.	
PR	30-AUG-1999;	99US-0151303P.	XX		
PR	31-AUG-1999;	99US-0151438P.	PN	EP1033405-A2.	
PR	01-SEP-1999;	99US-0151930P.	XX		
PR	07-SEP-1999;	99US-0152363P.	PD	06-SEP-2000.	
PR	10-SEP-1999;	99US-0153070P.	XX		
PR	13-SEP-1999;	99US-0153758P.	PF	25-FEB-2000;	2000EP-00301439.
PR	15-SEP-1999;	99US-0154018P.	XX		
PR	16-SEP-1999;	99US-0154039P.	PR	25-FEB-1999;	99US-0121825P.
PR	20-SEP-1999;	99US-0154779P.	PR	05-MAR-1999;	99US-0123180P.
PR	22-SEP-1999;	99US-0155139P.	PR	09-MAR-1999;	99US-0123548P.
PR	23-SEP-1999;	99US-0155486P.	PR	23-MAR-1999;	99US-0125788P.
PR	24-SEP-1999;	99US-0155659P.	PR	25-MAR-1999;	99US-0126264P.
PR	28-SEP-1999;	99US-0156458P.	PR	29-MAR-1999;	99US-0126785P.
PR	29-SEP-1999;	99US-0156596P.	PR	01-APR-1999;	99US-0127462P.
PR	04-OCT-1999;	99US-0157117P.	PR	06-APR-1999;	99US-0128234P.
PR	05-OCT-1999;	99US-0157753P.	PR	08-APR-1999;	99US-0128714P.
PR	06-OCT-1999;	99US-0157865P.	PR	16-APR-1999;	99US-0129845P.
PR	07-OCT-1999;	99US-0158029P.	PR	19-APR-1999;	99US-0130077P.
PR	08-OCT-1999;	99US-0158232P.	PR	21-APR-1999;	99US-0130449P.
PR	12-OCT-1999;	99US-0158369P.	PR	23-APR-1999;	99US-0130510P.
PR	13-OCT-1999;	99US-0159293P.	PR	28-APR-1999;	99US-0131449P.
PR	13-OCT-1999;	99US-0159294P.	PR	30-APR-1999;	99US-0132048P.
PR	13-OCT-1999;	99US-0159295P.	PR	30-APR-1999;	99US-0132407P.
PR	14-OCT-1999;	99US-0159329P.	PR	04-MAY-1999;	99US-0132484P.
PR	14-OCT-1999;	99US-0159330P.	PR	05-MAY-1999;	99US-0132485P.
PR	14-OCT-1999;	99US-0159331P.	PR	06-MAY-1999;	99US-0132486P.
PR	14-OCT-1999;	99US-0159637P.	PR	06-MAY-1999;	99US-0132487P.
PR	18-OCT-1999;	99US-0159638P.	PR	07-MAY-1999;	99US-0132863P.
PR	21-OCT-1999;	99US-0159584P.	PR	11-MAY-1999;	99US-0134256P.
PR	21-OCT-1999;	99US-0160741P.	PR	14-MAY-1999;	99US-0134218P.
PR	21-OCT-1999;	99US-0160767P.	PR	14-MAY-1999;	99US-0134219P.
PR	21-OCT-1999;	99US-0160768P.	PR	14-MAY-1999;	99US-0134221P.
PR	21-OCT-1999;	99US-0160770P.	PR	14-MAY-1999;	99US-0134370P.
PR	21-OCT-1999;	99US-0160814P.	PR	18-MAY-1999;	99US-0134768P.
PR	21-OCT-1999;	99US-0160815P.	PR	19-MAY-1999;	99US-0134941P.
PR	22-OCT-1999;	99US-0160980P.	PR	20-MAY-1999;	99US-0135124P.
PR	22-OCT-1999;	99US-0160981P.	PR	21-MAY-1999;	99US-0135353P.
PR	22-OCT-1999;	99US-0160989P.	PR	24-MAY-1999;	99US-0135629P.
PR	25-OCT-1999;	99US-0161404P.	PR	25-MAY-1999;	99US-0136021P.
PR	25-OCT-1999;	99US-0161405P.	PR	27-MAY-1999;	99US-0136392P.
PR	25-OCT-1999;	99US-0161406P.	PR	28-MAY-1999;	99US-0136782P.
PR	26-OCT-1999;	99US-0161359P.	PR	01-JUN-1999;	99US-0137222P.
PR	26-OCT-1999;	99US-0161360P.	PR	03-JUN-1999;	99US-0137528P.
PR	26-OCT-1999;	99US-0161361P.	PR	04-JUN-1999;	99US-0137502P.
PR	28-OCT-1999;	99US-0161920P.	PR	07-JUN-1999;	99US-0137724P.
PR	28-OCT-1999;	99US-0161992P.	PR	08-JUN-1999;	99US-0138094P.
PR	28-OCT-1999;	99US-0161993P.			

PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.

PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 100.0%; Score 45; DB 3; Length 340;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7

PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 100.0%; Score 45; DB 3; Length 398;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
Db 163 RPQPPPP 169

RESULT 17
ADN73283
ID ADN73283 standard; protein; 398 AA.

XX AC ADN73283;
XX DT 15-JUL-2004 (first entry)
XX DE Thale cress protein upregulated in E2Fa/Dpa expressing plants SeqID 1178.
XX KW plant; transgenic; E2Fa/Dpa transcription factor; growth regulator;
XX KW animal feed product; thale cress; cell wall biosynthesis;
XX KW nitrogen metabolism; carbon metabolism.
XX OS Arabidopsis thaliana.
XX PN WO2004035798-A2.
XX PD 29-APR-2004.
XX PF 20-OCT-2003; 2003WO-EP011658.
XX PR 18-OCT-2002; 2002EP-00079408.
XX PA (CROP-) CROPDESIGN NV.
XX PI Inze D, De Veylder L, Vlieghe K;
XX DR WPI; 2004-348466/32.
XX DR N-PSDB; ADN73282.
XX PT Altering plant characteristics, useful for producing plants for enzyme or
PT pharmaceutical production comprises modifying in a plant, expression of
PT one or more nucleic acids and/or modifying level or activity of one or
PT more proteins.
XX PS Claim 1; SEQ ID NO 1178; 134pp; English.
XX CC This invention relates to a novel method for altering one or more plant
CC characteristics. Specifically, it refers to identifying genes that are up
CC - or down-regulated in transgenic plants overexpressing the heterodimeric
CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to
CC alter plant characteristics accordingly. The present invention describes
CC generating transgenic plants for the production of growth regulators,
CC enzymes, therapeutics, pharmaceuticals and animal feed products, where
CC the altered plant characteristics are selected from increased yield or
CC biomass, enhanced survival capacity, stress tolerance, plant architecture
CC or physiology, altered endoreduplication, biochemistry, signal
CC transduction, storage lipid mobilisation and/or altered photosynthesis,
CC each relative to the corresponding wild type plants. Accordingly, these
CC sequences can also be useful as positive or negative selectable markers
CC during transformation of cells or tissues. The identified genes play a
CC role in a variety of biological processes such as DNA replication, cell
CC wall biosynthesis, nitrogen and/ or carbon metabolism or they function as
CC transcription factors. This polypeptide sequence is thale cress protein
CC expressed by a gene upregulated 1.3 fold or more in plants overexpressing
CC the E2Fa/Dpa transcription factor, given in an exemplification of the
XX SQ invention.
SQ Sequence 398 AA;

Query Match 100.0%; Score 45; DB 8; Length 398;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
Db 163 RPQPPPP 169

RESULT 18
AAG06369
ID AAG06369 standard; protein; 412 AA.
XX AC AAG06369;
XX

DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 3118.
DE
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.

PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.

PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 100.0%; Score 45; DB 3; Length 412;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 177 RPQPPPP 183

RESULT 19
AAW72182
ID AAW72182 standard; protein; 432 AA.

XX AAW72182;
XX
DT 13-JAN-1999 (first entry)
XX
DE HSV-2 strain SB5 Contig ID 15 ORF#15 protein.

XX
KW HSV-2 strain SB5; immunological response induction; therapy;
KW antiviral identification; viral protein inhibitor.

OS Herpes simplex virus 2.
XX
PN WO9820016-A1.
XX
PD 14-MAY-1998.
XX
PF 31-OCT-1997; 97WO-US020016.
XX
PR 04-NOV-1996; 96US-0030279P.
PR 09-JUN-1997; 97US-0049018P.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI Esser KM, Chan JY, Dabrowski-Amaral CE, Delvecchio AM, Dillon SB;
PI Leary JJ;
XX
DR WPI; 1998-286847/25.
DR N-PSDB; AAV62176.
XX
PT Herpes simplex virus type-2 sequences - useful in, e.g. prevention and
PT treatment of infection or inducing immunological response in mammal.
XX
PS Claim 10; Page 115; 748pp; English.
XX
CC This sequence represents a Herpes simplex virus type-2 (HSV-2) protein
CC sequence of the invention. This sequence was isolated from a HSV-2 strain
CC SB5 (deposited as ATCC VR-2546) DNA fragment designated Contig ID 15.
CC Based on homology, this sequence is a probable DNA packaging protein. The
CC proteins can be used for the treatment or prevention of disease, to
CC induce an immunological response in a mammal or to identify inhibitors,
CC activators or novel antivirals. Antagonists of the proteins can be used
CC to inhibit a viral polypeptide. The DNA sequence or a vector containing
CC it can also be used to induce an immunological response in a mammal
XX
SQ Sequence 432 AA;

Query Match 100.0%; Score 45; DB 2; Length 432;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 377 RPQPPPP 383

RESULT 20
ABB70063
ID ABB70063 standard; protein; 446 AA.

XX ABB70063;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 36981.

XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.

XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.

XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.

XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX

DR WPI; 2001-656860/75.
DR N-PSDB; ABL14166.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Disclosure; SEQ ID NO 36981; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 446 AA;

Query Match 100.0%; Score 45; DB 4; Length 446;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
Db 231 RPQPPPP 237

RESULT 21
AAW72030
ID AAW72030 standard; protein; 497 AA.
XX
AC AAW72030;
DT 07-DEC-1998 (first entry)
XX
DE HSV-2 strain SB5 Contig ID 103 ORF#8 protein.
XX
KW HSV-2 strain SB5; immunological response induction; therapy;
KW antiviral identification; viral protein inhibitor.
XX
OS Herpes simplex virus 2.
XX
PN WO9820016-A1.
XX
PD 14-MAY-1998.
XX
PF 31-OCT-1997; 97WO-US020016.
XX
PR 04-NOV-1996; 96US-0030279P.
PR 09-JUN-1997; 97US-0049018P.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI Esser KM, Chan JY, Dabrowski-Amaral CE, Delvecchio AM, Dillon SB;
PI Leary JJ;
XX
DR WPI; 1998-286847/25.
DR N-PSDB; AAV62133.
XX
PT Herpes simplex virus type-2 sequences - useful in, e.g. prevention and
PT treatment of infection or inducing immunological response in mammal.
XX
PS Claim 10; Page 50; 748pp; English.
XX
CC This sequence represents a Herpes simplex virus type-2 (HSV-2) protein
CC sequence of the invention. This sequence was isolated from a HSV-2 strain
CC SB5 (deposited as ATCC VR-2546) DNA fragment designated Contig ID 103.
CC Based on homology, this sequence is a 38K protein - human herpesvirus 1.
CC The proteins can be used for the treatment or prevention of disease, to

CC induce an immunological response in a mammal or to identify inhibitors,
CC activators or novel antivirals. Antagonists of the proteins can be used
CC to inhibit a viral polypeptide. The DNA sequence or a vector containing
CC it can also be used to induce an immunological response in a mammal
XX
SQ Sequence 497 AA;

Query Match 100.0%; Score 45; DB 2; Length 497;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
Db 442 RPQPPPP 448

RESULT 22
AAW72132
ID AAW72132 standard; protein; 497 AA.
XX
AC AAW72132;
XX
DT 23-DEC-1998 (first entry)
XX
DE HSV-2 strain SB5 Contig ID 16 ORF#8 protein.
XX
KW HSV-2 strain SB5; immunological response induction; therapy;
KW antiviral identification; viral protein inhibitor.
XX
OS Herpes simplex virus 2.
XX
PN WO9820016-A1.
XX
PD 14-MAY-1998.
XX
PF 31-OCT-1997; 97WO-US020016.
XX
PR 04-NOV-1996; 96US-0030279P.
PR 09-JUN-1997; 97US-0049018P.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI Esser KM, Chan JY, Dabrowski-Amaral CE, Delvecchio AM, Dillon SB;
PI Leary JJ;
XX
DR WPI; 1998-286847/25.
DR N-PSDB; AAV62160.
XX
PT Herpes simplex virus type-2 sequences - useful in, e.g. prevention and
PT treatment of infection or inducing immunological response in mammal.
XX
PS Claim 10; Page 95; 748pp; English.
XX
CC This sequence represents a Herpes simplex virus type-2 (HSV-2) protein
CC sequence of the invention. This sequence was isolated from a HSV-2 strain
CC SB5 (deposited as ATCC VR-2546) DNA fragment designated Contig ID 16.
CC Based on homology, this sequence is a probable DNA packaging protein. The
CC proteins can be used for the treatment or prevention of disease, to
CC induce an immunological response in a mammal or to identify inhibitors,
CC activators or novel antivirals. Antagonists of the proteins can be used
CC to inhibit a viral polypeptide. The DNA sequence or a vector containing
CC it can also be used to induce an immunological response in a mammal
XX
SQ Sequence 497 AA;

Query Match 100.0%; Score 45; DB 2; Length 497;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
Db 442 RPQPPPP 448

```
RESULT 23
AAY29515
ID AAY29515 standard; protein; 527 AA.
XX
AC AAY29515;
XX
DT 13-OCT-1999 (first entry)
XX
DE Human lung tumour protein SAL-50 2nd predicted amino acid sequence.
XX
KW Human; lung tumour protein; therapy; diagnosis; lung cancer; vaccine;
KW immunotherapy; detection; inhibition.
XX
OS Homo sapiens.
XX
PN WO9938973-A2.
XX
PD 05-AUG-1999.
XX
PF 26-JAN-1999; 99WO-US001642.
XX
PR 28-JAN-1998; 98US-00015022.
PR 28-JAN-1998; 98US-00015029.
PR 18-MAR-1998; 98US-00040828.
PR 18-MAR-1998; 98US-00040831.
PR 23-JUL-1998; 98US-00122191.
PR 23-JUL-1998; 98US-00122192.
PR 22-DEC-1998; 98US-00219245.
XX
PA (CORI-) CORIXA CORP.
XX
PI Reed SG, Lodes MJ, Frudakis TN, Mohamath R;
XX
DR WPI; 1999-479187/40.
DR N-PSDB; AAZ07194.
XX
Lung tumor specific polynucleotides for inhibiting the development of
lung cancer.
XX
PS Claim 3; Page 170-171; 171pp; English.
XX
The present invention describes lung tumour specific polynucleotides and
tumour antigens. AAZ07144 to AAZ07246 and AAZ08301 to AAZ08325 represent
specifically claimed polynucleotides, and AAY29486 to AAY29571 represent
amino acid sequences from the present invention. The lung tumour specific
polynucleotides and polypeptides can be used in pharmaceutical
compositions and vaccines to inhibit the development of lung cancer. They
can also be used to detect lung cancer in a patient. Probes and
antibodies derived from the lung tumour sequences are useful in detection
of lung cancer
XX
SQ Sequence 527 AA;

Query Match 100.0%; Score 45; DB 2; Length 527;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
Db 209 RPQPPPP 215

RESULT 24
AAB44492
ID AAB44492 standard; protein; 527 AA.
XX
AC AAB44492;
XX
DT 05-FEB-2001 (first entry)
XX
DE Human lung tumour-specific antigen encoded by cDNA #100.
XX
```

```
KW Lung tumour protein; lung cancer; cytostatic; vaccine.
XX
OS Homo sapiens.
XX
PN WO200060077-A2.
XX
PD 12-OCT-2000.
XX
PF 30-MAR-2000; 2000WO-US008560.
XX
PR 02-APR-1999; 99US-00285323.
PR 09-AUG-1999; 99US-00370838.
PR 30-DEC-1999; 99US-00476235.
PR 03-MAR-2000; 2000US-00518809.
XX
PA (CORI-) CORIXA CORP.
XX
PI Reed SG, Lodes MJ, Mohamath R, Secríst H;
XX
DR WPI; 2000-638466/61.
DR N-PSDB; AAC79147.
XX
Novel lung tumor polypeptides and polynucleotides, useful for detecting,
monitoring or treating cancer, especially lung cancer.
XX
PS Example 4; Page 185-186; 243pp; English.
XX
The present sequence is given in a specification relating to compounds
for therapy and diagnosis of lung cancer. Polypeptides comprising at
least an immunogenic part of a lung tumour protein are disclosed. The
polypeptides are useful for inhibiting the development of cancer,
especially lung cancer. Samples of T cells expressing the polypeptides
may be used to inhibit the development of cancer. The polypeptides are
also useful for detecting and monitoring the progression of cancer,
especially lung cancer
XX
SQ Sequence 527 AA;

Query Match 100.0%; Score 45; DB 3; Length 527;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
Db 209 RPQPPPP 215

RESULT 25
AAE13805
ID AAE13805 standard; protein; 527 AA.
XX
AC AAE13805;
XX
DT 26-FEB-2002 (first entry)
XX
DE Human lung tumour-specific protein SAL-50 #2.
XX
KW Human; lung tumour protein; immunostimulant; cytostatic; gene therapy;
KW antisense-therapy; vaccine; immune response; lung cancer; SAL-50.
XX
OS Homo sapiens.
XX
PN WO200172295-A2.
XX
PD 04-OCT-2001.
XX
PF 28-MAR-2001; 2001WO-US009991.
XX
PR 29-MAR-2000; 2000US-00538037.
PR 05-JUN-2000; 2000US-00588937.
PR 18-AUG-2000; 2000US-00640878.
PR 22-SEP-2000; 2000US-0234517P.
PR 01-NOV-2000; 2000US-00704512.
XX
```

PR 14-DEC-2000; 2000US-00738973.
XX (CORI-) CORIXA CORP.
PA Reed SG, Lodes MJ, Mohamath R, Secrist H, Benson DR, Indirias CY;
XX Henderson RA, Fling SP, Algate PA, Elliot M, Mannion J, Kalos MD;
PI WPI; 2001-639201/73.
XX N-PSDB; AAD23222.
DR New human lung-specific polynucleotides and polypeptides for the
XX diagnosis and treatment of disease e.g. lung cancer.
DR Example 4; Page 237-238; 378pp; English.
PS The invention relates to isolated lung tumour-specific proteins and their
XX corresponding cDNA molecules. Lung tumour-specific proteins and their
CC antigen-presenting cells are useful for stimulating and/or expanding T
CC cells specific for a tumour protein, and for inhibiting the development
CC of cancer. The invention also relates to a composition useful for
CC stimulating an immune response, and for treating cancer. The lung tumour
CC specific oligonucleotide is useful in gene therapy and for diagnosis,
CC detection and treatment of lung cancer. The present sequence is human
CC lung tumour-specific protein
XX
SQ Sequence 527 AA;
Query Match 100.0%; Score 45; DB 4; Length 527;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RPQPPPP 7
Db 209 RPQPPPP 215
RESULT 26
ADD66524
ID ADD66524 standard; protein; 527 AA.
XX
AC ADD66524;
DT 15-JAN-2004 (first entry)
XX
DE Human lung tumour-specific related protein, SEQ ID No 216.
XX expression control; cancer; T cell; tumour; immune; cytostatic; vaccine;
KW human; lung tumour-specific.
XX
OS Homo sapiens.
XX
PN WO200292001-A2.
PD 21-NOV-2002.
XX
PF 10-MAY-2002; 2002WO-US014975.
XX
PR 11-MAY-2001; 2001US-00854133.
XX (CORI-) CORIXA CORP.
PA Lodes MJ, Wang T, Fan L, Algate PA, Mcneill PD;
XX WPI; 2003-120592/11.
DR New polynucleotide and polypeptide, useful for preparing a composition
XX for diagnosing, treating or preventing cancer.
PS Example 4; SEQ ID NO 216; 494pp; English.
XX The invention relates to a novel isolated polynucleotide comprising one
CC of 32 47-6080 base pair sequences, given in the specification, or their
CC complements or degenerate variants, at least 20 contiguous residues of a

CC sequence in, or having at least 75 or 90 % identity with the isolated
CC polynucleotide, or that hybridise with the polynucleotide. The invention
CC further comprises: an isolated polypeptide; an expression vector
CC comprising the polynucleotide operably linked to an expression control
CC sequence; a host cell transformed or transfected with the expression
CC vector; an isolated antibody or its antigen-binding fragment that
CC specifically binds to the polypeptide; a method for detecting the
CC presence of a cancer in a patient; a fusion protein comprising the
CC polypeptide; an oligonucleotide that hybridises to the isolated
CC polynucleotide under moderately stringent conditions; a method for
CC stimulating and/or expanding T cells specific for a tumour protein; an
CC isolated T cell population; a composition comprising a first component
CC consisting of carriers and immunostimulants and a second component; a
CC method for stimulating an immune response in a patient; a method for
CC treating cancer in a patient; a method for determining cancer in a
CC patient; a diagnostic kit comprising at least one oligonucleotide or
CC antibody and a detection reagent comprising a reporter group; and a
CC method for inhibiting the development of cancer in a patient. The
CC compositions of the invention have cytostatic activity and can be used to
CC create a vaccine. The isolated polynucleotide is useful for preparing a
CC composition for diagnosing, treating or preventing cancer. This sequence
CC represents a human lung tumour-specific protein relating to the
CC invention.
XX
SQ Sequence 527 AA;
Query Match 100.0%; Score 45; DB 7; Length 527;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RPQPPPP 7
Db 209 RPQPPPP 215
RESULT 27
ADE87778
ID ADE87778 standard; protein; 527 AA.
XX
AC ADE87778;
DT 29-JAN-2004 (first entry)
XX
DE Human lung tumour antigen polypeptide #86.
XX Human; lung tumour antigen; cancer; lung cancer; CD4+; T cell;
KW immune response; immunostimulant; cytostatic.
XX
OS Homo sapiens.
XX
PN US2003118599-A1.
PD 26-JUN-2003.
XX
PF 10-MAY-2002; 2002US-00144649.
XX
PR 02-APR-1999; 99US-00285323.
PR 09-AUG-1999; 99US-00370838.
PR 30-DEC-1999; 99US-00476235.
PR 03-MAR-2000; 2000US-00518809.
PR 29-MAR-2000; 2000US-00538037.
PR 05-JUN-2000; 2000US-00588937.
PR 18-AUG-2000; 2000US-00640878.
PR 20-SEP-2000; 2000US-00667170.
PR 01-NOV-2000; 2000US-00704512.
PR 14-DEC-2000; 2000US-00738973.
PR 11-MAY-2001; 2001US-00854133.
XX (CORI-) CORIXA CORP.
PA Algate PA, Lodes MJ, Wang T, Fan L, Mcneill PD;
XX WPI; 2003-897103/82.
DR

DR N-PSDB; ADB87743.
XX
PT New polynucleotides encode lung tumor antigens and are useful to
PT stimulate an immune response or detect or treat a cancer in a patient,
PT particularly lung cancer.
XX
PS Example 4; SEQ ID NO 216; 63pp; English.
XX
CC The invention relates to polynucleotides encoding lung tumour antigens.
CC The invention also relates to the polypeptides encoded by the
CC polynucleotides, isolated antibodies or antigen-binding fragments that
CC specifically bind the polypeptides and a method for detecting cancer in a
CC patient, comprising obtaining a biological sample from the patient,
CC contacting the sample with a binding agent that binds a polypeptide of
CC the invention, detecting in the sample an amount of polypeptide that
CC binds to the binding agent, and comparing the amount of polypeptide to a
CC predetermined cut-off value. T cells specific for a tumour protein can be
CC stimulated and/or expanded by contacting the T cells with a polypeptide,
CC polynucleotide or an antigen-presenting cell that expresses a
CC polynucleotide. Cancer development can be inhibited by incubating CD4+
CC and/or CD8+ T cells isolated from a patient with a polypeptide,
CC polynucleotide or an antigen-presenting cell that expresses a
CC polypeptide, so that the T cells proliferate. The invention is used to
CC stimulate an immune response or to detect or treat a cancer in a patient,
CC particularly lung cancer. This sequence represents a human lung tumour
CC antigen polypeptide of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification but was obtained in
CC electronic format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 527 AA;

Query Match 100.0%; Score 45; DB 7; Length 527;
Best Local Similarity 100.0%; Pred. NO. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
|||
Db 209 RPQPPPP 215

RESULT 28
AAU10545
ID AAU10545 standard; protein; 668 AA.

XX AC AAU10545;

XX DT 14-FEB-2002 (first entry)

XX DE Rat synapsin 1B (YSG8) polypeptide.

XX YSG; YSG8; schizophrenia; chronic animal model; LCGU; netrin receptor;
KW local cerebral glucose utilisation; phosphodiesterase 1-alpha; UNC5H1;
KW calcium-independent alpha-latrotoxin receptor; CIRL; trkE; synapsin 1A;
KW epithelial discoidin domain receptor 1; synapsin 1B; neuroleptic;
KW tumour necrosis factor alpha; TNF-alpha; rat.

XX OS Rattus sp.

XX PN WO200175440-A2.

XX PD 11-OCT-2001.

XX PF 02-APR-2001; 2001WO-GB001486.

XX PR 31-MAR-2000; 2000GB-00007880.

XX PR 26-MAY-2000; 2000GB-00012768.

XX PA (WELF-) WELFIDE CORP.

XX PI Cochran S, Paterson G, Ohashi Y, Morris B, Pratt J;

XX WPI; 2002-010813/01.

XX N-PSDB; AAS16845.

XX Novel chronic animal model of schizophrenia, useful for identifying anti-
PT psychotic drugs and genes that are associated with schizophrenia.
XX Claim 1; Fig 9d; 79pp; English.

XX The invention relates to YSG polynucleotide fragments for use in
CC diagnosing and/or developing treatments for schizophrenia using chronic
CC animal models. The polynucleotides and their encoded polypeptides are
CC used for identification of compounds which modulate the expression of YSG
CC molecules, leading to the manufacture of schizophrenia medications. The
CC sequences can also be used for testing candidate compounds for any effect
CC on the polypeptides. Anti-schizophrenic effects of a compound can be
CC determined by measuring local cerebral glucose utilisation (LCGU) or
CC comparing its expression level with that of a control group. The
CC sequences are useful in the identification of genes associated with
CC schizophrenic states and in the development of an antibody. The sequences
CC of the invention include phosphodiesterase 1-alpha, calcium-independent
CC alpha-latrotoxin receptors (CIRL)-1,2&3, epithelial discoidin domain
CC receptor 1 (trkE), netrin receptor (UNC5H1), synapsins 1A and AB and
CC tumour necrosis factor (TNF) alpha. This sequence represents rat synapsin
CC 1B (YSG8) polypeptide
XX Sequence 668 AA;

Query Match 100.0%; Score 45; DB 5; Length 668;
Best Local Similarity 100.0%; Pred. NO. 4.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
|||
Db 27 RPQPPPP 33

RESULT 29
ADX05987
ID ADX05987 standard; protein; 669 AA.

XX AC ADX05987;

XX DT 21-APR-2005 (first entry)

XX DE Cyclin-dependent kinase modulation biomarker SEQ ID NO 552.

XX KW cytosstatic; cyclin-dependent kinase; cdk; biomarker.

XX OS Homo sapiens.

XX PN WO2005012875-A2.

XX PD 10-FEB-2005.

XX PF 29-JUL-2004; 2004WO-US024424.

XX PR 29-JUL-2003; 2003US-0490890P.

XX PA (BRIM) BRISTOL-MYERS SQUIBB CO.

XX PI Li M, Rupnow BA, Webster KR, Jackson DG, Wong TW;

XX DR WPI; 2005-163068/17.

XX DR N-PSDB; ADX05986.

XX Biomarkers useful for predicting or determining the response of a mammal
PT to a cancer treatment comprising administration of a modulator of cyclin-
PT dependent kinase activity.

XX Claim 5; SEQ ID NO 552; 141pp; English.

XX This invention describes a novel method of predicting or determining
CC whether a mammal will respond or is responding to an anti-cancer agent
CC that modulates cyclin-dependent kinase (cdk) activity. The method
CC comprises measuring the level of one or more biomarkers selected from

CC 2774 biomarkers given in the specification (nucleotide sequence SEQ ID
CC NO:1246 (Genbank EST W28729) is especially preferred). The method of the
CC invention is utilized in a kit for determining or predicting whether
CC patient would be susceptible or resistant to treatment by an agent
CC modulating cdk activity. The invention also describes a method for
CC utilizing individualized genetic profiles for treating diseases and
CC disorders based on patient's response and molecular level, specialized
CC microarrays comprising the biomarkers described, antibodies directed
CC against the biomarkers and a cell culture model to identify biomarkers.
CC The cdk modulator is preferably N-5-[(5-(1,1-Dimethylethyl)-2-
CC oxazolyl)methyl]thio]-2-thiazolyl-4-piperidine carboxamide, 0.5-L-
CC tartaric acid salt. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences. This
CC sequence represents a biomarker used in the method of the invention.
XX
SQ Sequence 669 AA;

Query Match 100.0%; Score 45; DB 9; Length 669;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 27 RPQPPPP 33

RESULT 30
AAU10544
ID AAU10544 standard; protein; 704 AA.
XX
AC AAU10544;
XX
DT 14-FEB-2002 (first entry)
XX
DE Rat synapsin 1A (YSG8) polypeptide.
XX
KW YSG; YSG8; schizophrenia; chronic animal model; LCGU; netrin receptor;
KW local cerebral glucose utilisation; phosphodiesterase 1-alpha; UNC5H1;
KW calcium-independent alpha-latrotoxin receptor; CIRL; trkE; synapsin 1A;
KW epithelial discoidin domain receptor 1; synapsin 1B; neuroleptic;
KW tumour necrosis factor alpha; TNF-alpha; rat.
XX
OS Rattus sp.
XX
PN WO200175440-A2.
XX
PD 11-OCT-2001.
XX
PF 02-APR-2001; 2001WO-GB001486.
XX
PR 31-MAR-2000; 2000GB-00007880.
PR 26-MAY-2000; 2000GB-00012768.
XX
PA (WELF-) WELFIDE CORP.
XX
PI Cochran S, Paterson G, Ohashi Y, Morris B, Pratt J;
XX
DR WPI; 2002-010813/01.
DR N-PSDB; AAS16844.
XX
PT Novel chronic animal model of schizophrenia, useful for identifying anti-
PT psychotic drugs and genes that are associated with schizophrenia.
XX
PS Disclosure; Fig 9b; 79pp; English.
XX
CC The invention relates to YSG polynucleotide fragments for use in
CC diagnosing and/or developing treatments for schizophrenia using chronic
CC animal models. The polynucleotides and their encoded polypeptides are
CC used for identification of compounds which modulate the expression of YSG
CC molecules, leading to the manufacture of schizophrenia medicaments. The
CC sequences can also be used for testing candidate compounds for any effect
CC on the polypeptides. Anti-schizophrenic effects of a compound can be

CC determined by measuring local cerebral glucose utilisation (LCGU) or
CC comparing its expression level with that of a control group. The
CC sequences are useful in the identification of genes associated with
CC schizophrenic states and in the development of an antibody. The sequences
CC of the invention include phosphodiesterase 1-alpha, calcium-independent
CC alpha-latrotoxin receptors (CIRL)-1,2&3, epithelial discoidin domain
CC receptor 1 (trkE), netrin receptor (UNC5H1), synapsins 1A and AB and
CC tumour necrosis factor (TNF) alpha. This sequence represents rat synapsin
CC 1A (YSG8) polypeptide
XX
SQ Sequence 704 AA;

Query Match 100.0%; Score 45; DB 5; Length 704;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 27 RPQPPPP 33

RESULT 31
ADE57712
ID ADE57712 standard; protein; 704 AA.
XX
AC ADE57712;
XX
DT 29-JAN-2004 (first entry)
XX
DE Rat Protein P09951, SEQ ID NO 3575.
XX
KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS Rattus norvegicus.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR WPI; 2003-268312/26.
DR GENBANK; P09951.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a

CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 704 AA;
SQ Query Match 100.0%; Score 45; DB 7; Length 704;
Best Local Similarity 100.0%; Pred. NO. 4.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RPQPPPP 7
Db 27 RPQPPPP 33

RESULT 32
AAM39328
ID AAM39328 standard; protein; 705 AA.

XX AC AAM39328;

XX DT 22-OCT-2001 (first entry)

XX DE Human polypeptide SEQ ID NO 2473.

XX KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.

XX OS Homo sapiens.

XX PN WO200153312-A1.

XX PD 26-JUL-2001.

XX PF 26-DEC-2000; 2000WO-US034263.

XX PR 23-DEC-1999; 99US-00471275.

XX PR 21-JAN-2000; 2000US-00488725.

XX PR 25-APR-2000; 2000US-00552317.

XX PR 20-JUN-2000; 2000US-00598042.

XX PR 19-JUL-2000; 2000US-00620312.

XX PR 03-AUG-2000; 2000US-00653450.

XX PR 14-SEP-2000; 2000US-00662191.

XX PR 19-OCT-2000; 2000US-00693036.

XX PR 29-NOV-2000; 2000US-00727344.

XX N-PSDB; AAI58484.

XX Novel nucleic acids and polypeptides, useful for treating disorders such

XX as central nervous system injuries.

XX Example 4; SEQ ID NO 2473; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification

XX SQ Sequence 705 AA;
SQ Query Match 100.0%; Score 45; DB 4; Length 705;
Best Local Similarity 100.0%; Pred. NO. 4.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RPQPPPP 7
Db 27 RPQPPPP 33

RESULT 33
ADX05985

XX ID ADX05985 standard; protein; 705 AA.

XX AC ADX05985;

XX DT 21-APR-2005 (first entry)

XX DE Cyclin-dependent kinase modulation biomarker SEQ ID NO 550.

XX KW cytostatic; cyclin-dependent kinase; cdk; biomarker.

XX OS Homo sapiens.

XX PN WO2005012875-A2.

XX PD 10-FEB-2005.

XX PF 29-JUL-2004; 2004WO-US024424.

XX PR 29-JUL-2003; 2003US-0490890P.

XX (BRIM) BRISTOL-MYERS SQUIBB CO.

XX Li M, Rupnow BA, Webster KR, Jackson DG, Wong TW;

XX WPI; 2005-163068/17.

XX N-PSDB; ADX05984.

XX Biomarkers useful for predicting or determining the response of a mammal
PT to a cancer treatment comprising administration of a modulator of cyclin-
PT dependent kinase activity.

XX Claim 5; SEQ ID NO 550; 141pp; English.

XX This invention describes a novel method of predicting or determining
CC whether a mammal will respond or is responding to an anti-cancer agent
CC that modulates cyclin-dependent kinase (cdk) activity. The method
CC comprises measuring the level of one or more biomarkers selected from
CC 2774 biomarkers given in the specification (nucleotide sequence SEQ ID
CC NO:1246 (Genbank EST W28729) is especially preferred). The method of the
CC invention is utilized in a kit for determining or predicting whether
CC patient would be susceptible or resistant to treatment by an agent
CC modulating cdk activity. The invention also describes a method for
CC utilizing individualized genetic profiles for treating diseases and

CC disorders based on patient's response and molecular level, specialized
CC microarrays comprising the biomarkers described, antibodies directed
CC against the biomarkers and a cell culture model to identify biomarkers.
CC The cdk modulator is preferably N-5-[(5-(1,1-Dimethylethyl)-2-
CC oxazolyl)methyl]thio]-2-thiazolyl-4-piperidine carboxamide, 0.5-L-
CC tartaric acid salt. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences. This
CC sequence represents a biomarker used in the method of the invention.

XX Sequence 705 AA;

SQ Query Match 100.0%; Score 45; DB 9; Length 705;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
Db 27 RPQPPPP 33

RESULT 34
AAM41114
ID AAM41114 standard; protein; 714 AA.

XX AAM41114;

DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 6045.

XX Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; Chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.

OS Homo sapiens.

XX WO200153312-A1.

PN 26-JUL-2001.

PD 26-DEC-2000; 2000WO-US034263.

XX 23-DEC-1999; 99US-00471275.

PR 21-JAN-2000; 2000US-00488725.

PR 25-APR-2000; 2000US-00552317.

PR 20-JUN-2000; 2000US-00598042.

PR 19-JUL-2000; 2000US-00620312.

PR 03-AUG-2000; 2000US-00653450.

PR 14-SEP-2000; 2000US-00662191.

PR 19-OCT-2000; 2000US-00693036.

XX 29-NOV-2000; 2000US-00727344.

PA (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;

XX WPI; 2001-442253/47.

DR N-PSDB; AAI60270.

XX Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.

XX Example 2; SEQ ID NO 6045; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with neurotropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful

CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification

XX Sequence 714 AA;

SQ Query Match 100.0%; Score 45; DB 4; Length 714;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
Db 35 RPQPPPP 41

RESULT 35
ABG26213
ID ABG26213 standard; protein; 735 AA.

XX ABG26213;

DT 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #26204.

DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

PN 11-OCT-2001.

PD 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

PI WPI; 2001-639362/73.

DR N-PSDB; AAS90400.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

XX Claim 20; SEQ ID NO 56572; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders

CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

SQ Sequence 735 AA;

Query Match 100.0%; Score 45; DB 4; Length 735;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7

Db 9 RPQPPPP 15

RESULT 36
ABG30148
ID ABG30148 standard; protein; 812 AA.

XX AC ABG30148;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #30139.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS94335.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.

XX PS Claim 20; SEQ ID NO 60507; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX CC sequences. (I) is useful as hybridisation probes, polymerase chain
XX CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX CC and in recombinant production of (II). The polynucleotides are also used
XX CC in diagnostics as expressed sequence tags for identifying expressed
XX CC genes. (I) is useful in gene therapy techniques to restore normal
XX CC activity of (II) or to treat disease states involving (II). (II) is
XX CC useful for generating antibodies against it, detecting or quantitating a
XX CC polypeptide in tissue, as molecular weight markers and as a food
XX CC supplement. (II) and its binding partners are useful in medical imaging
XX CC of sites expressing (II). (I) and (II) are useful for treating disorders
XX CC involving aberrant protein expression or biological activity. The
XX CC polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

SQ Sequence 812 AA;

Query Match 100.0%; Score 45; DB 4; Length 812;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7

Db 445 RPQPPPP 451

RESULT 37
ABG02079
ID ABG02079 standard; protein; 874 AA.

XX AC ABG02079;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #2070.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS66266.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.

XX PS Claim 20; SEQ ID NO 32438; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX CC sequences. (I) is useful as hybridisation probes, polymerase chain
XX CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX CC and in recombinant production of (II). The polynucleotides are also used
XX CC in diagnostics as expressed sequence tags for identifying expressed
XX CC genes. (I) is useful in gene therapy techniques to restore normal
XX CC activity of (II) or to treat disease states involving (II). (II) is
XX CC useful for generating antibodies against it, detecting or quantitating a
XX CC polypeptide in tissue, as molecular weight markers and as a food
XX CC supplement. (II) and its binding partners are useful in medical imaging
XX CC of sites expressing (II). (I) and (II) are useful for treating disorders
XX CC involving aberrant protein expression or biological activity. The
XX CC polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 874 AA;

Query Match 100.0%; Score 45; DB 4; Length 874;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
Db 445 RPQPPPP 451

RESULT 38
ABG00401
ID ABG00401 standard; protein; 1074 AA.
XX
AC ABG00401;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #392.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS64588.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 30760; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic

CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1074 AA;

Query Match 100.0%; Score 45; DB 4; Length 1074;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
Db 445 RPQPPPP 451

RESULT 39
AAE21729
ID AAE21729 standard; protein; 1413 AA.
XX
AC AAE21729;
XX
DT 16-JUL-2002 (first entry)
XX
DE Human PKIN-24 protein.
XX
KW Human; kinase; enzyme; PKIN-24 protein; immune system disorder; anaemia;
KW acquired immune deficiency syndrome; thymic hypoplasia; Crohn's disease;
KW asthma; neurological disorder; epilepsy; Charcot-Marie-Tooth disease;
KW AIDS; seizures; cell proliferative disorder; cancer; adenocarcinoma;
KW leukaemia; lymphoma; melanoma; myeloma; sarcoma; developmental disorder;
KW Down's syndrome; gene therapy; protein therapy; cytostatic.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1. .20
FT /label= Signal_peptide
FT Protein 21. .1413
FT /label= Human_mature_PKIN-19_protein
FT Domain 133. .404
FT /note= "Eukaryotic protein kinase domain"
FT Domain 136. .400
FT /note= "Protein kinase domain"
FT Domain 137. .400
FT /note= "Protein kinase domain"
FT Domain 137. .398
FT /note= "Protein kinase domain"
FT Domain 210. .223
FT /note= "Tyrosine kinase catalytic domain"
FT Domain 256. .274
FT /note= "Tyrosine kinase catalytic domain"
FT Domain 305. .315
FT /note= "Tyrosine kinase catalytic domain"
FT Domain 331. .353
FT /note= "Tyrosine kinase catalytic domain"
FT Domain 380. .402
FT /note= "Tyrosine kinase catalytic domain"
XX WO200218557-A2.
PN 07-MAR-2002.
XX
XX 31-AUG-2001; 2001WO-US027219.
PF
XX 31-AUG-2000; 2000US-0229873P.
PR 08-SEP-2000; 2000US-0231357P.
PR 14-SEP-2000; 2000US-0232654P.
PR 22-SEP-2000; 2000US-0234902P.
PR 29-SEP-2000; 2000US-0236499P.
PR 06-OCT-2000; 2000US-0238389P.
PR 13-OCT-2000; 2000US-0240542P.
XX

PA (INCY-) INCYTE GENOMICS INC.
XX
PI Bandman O, Nguyen DB, Walia NK, Hafalia AJA, Yao MG, Gandhi AR;
PI Gururajan R, Ding L, Patterson C, Yue H, Baughn MR, Tribouley CM;
PI Thornton M, Elliott VS, Lu Y, Ison CH, Au-Young J, Tang YT;
PI Azimzai Y, Burrill JD, Marcus GA, Zingler KA, Lu DAM, Lal PG;
PI Ramkumar J, Warren BA, Kearney L, Policky JL, Thangavelu K;
PI Burford N;
XX
DR WPI; 2002-329769/36.
DR N-PSDB; AAD34321.
XX
PT New human kinases, useful for diagnosing, treating or preventing immune
PT system disorders (e.g. Crohn's disease), neurological disorders (e.g.
PT epilepsy), or cell proliferative disorders (e.g. cancers such as leukemia
PT or lymphoma).
XX
PS Claim 79; Page 189-193; 218pp; English.
XX
CC The present invention relates to human kinases (PKIN) and polynucleotides
CC encoding such proteins. PKIN sequences of the invention are useful for
CC diagnosing, treating or preventing disorders associated with aberrant
CC expression of PKIN, particularly immune system disorders (e.g. acquired
CC immune deficiency syndrome (AIDS), thymic hypoplasia, Crohn's disease,
CC anaemia, asthma), neurological disorders (e.g. epilepsy, Charcot-Marie-
CC Tooth disease or seizures), cell proliferative disorders (e.g. cancers
CC such as adenocarcinoma, leukaemia, lymphoma, melanoma, sarcoma),
CC and developmental disorders (e.g. Down's syndrome). They are also used in
CC gene therapy and protein therapy. The present sequence is human PKIN-24
CC protein
XX
SQ Sequence 1413 AA;

Query Match 100.0%; Score 45; DB 5; Length 1413;
Best Local Similarity 100.0%; Pred. No. 8.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 1095 RPQPPPP 1101

RESULT 40
ADC59465
ID ADC59465 standard; protein; 1460 AA.
XX
AC ADC59465;
XX
DT 18-DEC-2003 (first entry)
XX
DE Novel human tyrosine kinase protein.
XX
KW Enzyme; Tyrosine kinase; antiinflammatory; cytostatic;
KW antiarteriosclerotic; immunomodulatory; neuroprotective; gene therapy;
KW cell proliferation; Human.
XX
OS Homo sapiens.
XX
PN JP2003024075-A.
XX
PD 28-JAN-2003.
XX
PF 11-JUL-2001; 2001JP-00211135.
XX
PR 11-JUL-2001; 2001JP-00211135.
XX
PA (KAZU-) ZH KAZUSA DNA KENKYUSHO.
XX
DR WPI; 2003-545691/52.
DR N-PSDB; ADC59464.
XX
PT Novel recombinant protein useful for controlling proliferation of various
PT cells, has biological activity substantially homogeneous as tyrosine

PT kinase.
XX
PS Disclosure; SEQ ID NO 1; 22pp; Japanese.
XX
CC This invention relates to a novel recombinant protein which has tyrosine
CC kinase activity. The invention also discloses a pharmaceutical
CC composition comprising a DNA and protein sequence of the invention and an
CC antibody which binds specifically to the tyrosine kinase protein. The
CC protein of the invention may have antiinflammatory, cytostatic,
CC antiarteriosclerotic, immunomodulatory and neuroprotective activity and
CC may be used in gene therapy. The sequences of the invention are useful
CC for controlling proliferation of various cells, differentiation and
CC signal transduction. The protein is useful as a reagent for screening the
CC compound or its salt which inhibits the biological activity of the
CC tyrosine kinase. The sequences of the invention are also useful in the
CC diagnosis, prevention and treatment of diseases such as immune disorders,
CC neurological diseases, inflammation, cancer and arteriosclerosis. The
CC nucleic acid sequence of the invention may also useful as a detection
CC probe. The present sequence represents the novel human tyrosine kinase
CC protein of the invention.
XX
SQ Sequence 1460 AA;

Query Match 100.0%; Score 45; DB 7; Length 1460;
Best Local Similarity 100.0%; Pred. No. 8.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 1142 RPQPPPP 1148

RESULT 41
ADU04632
ID ADU04632 standard; protein; 1460 AA.
XX
AC ADU04632;
XX
DT 13-JAN-2005 (first entry)
XX
DE Human KIAA 1883 polypeptide.
XX
KW Human; KIAA 1883; lung cancer; marker; diagnosis.
XX
OS Homo sapiens.
XX
PN WO2004091511-A2.
XX
PD 28-OCT-2004.
XX
PF 12-APR-2004; 2004WO-US011193.
XX
PR 10-APR-2003; 2003US-0462028P.
XX
PA (GENZ) GENZYME CORP.
XX
PI Roberts BL;
XX
DR WPI; 2004-766692/75.
DR N-PSDB; ADU04631.
XX
PT Detecting neoplasia in lung cells comprises detecting the level of
PT expression of at least one gene selected from EGFR-RS, RYK, TNFRS25,
PT TRPM7, UNC5H2, KCP3 and KIAA 1883.
XX
PS Disclosure; SEQ ID NO 14; 80pp; English.
XX
CC The present sequence is that of human KIAA 1833 polypeptide. KIAA 1833
CC cDNA was identified in a sequencing project for large proteins of
CC unidentified genes. The invention provides a method of diagnosing the
CC neoplastic state of a lung cell by screening for the presence of a
CC differentially expressed gene. The gene is selected from EGFR-RS, RYK,
CC TNFRSF25, TRPM7, KCP3 and KIAA 1833, which are expressed at higher levels

CC in neoplastic lung cell or lung tumour cells than in healthy lung cells,
CC or UNC5H2, which is expressed at a higher level in healthy lung cells
CC than in neoplastic lung cells. The genes and their products provide
CC diagnostic and prognostic markers and therapeutic targets. Databases
CC containing quantitative full or partial transcripts or protein sequences
CC isolated from a cell sample can be searched and analysed for the presence
CC and amount of transcript or expressed gene product. The methods are
CC particularly useful for aiding in the diagnosis of non-small cell lung
CC cancer. The invention also provides methods of screening for potential
CC therapeutic agents for the reversal of the neoplastic condition. Claimed
CC methods for inhibiting the growth of a neoplastic lung cell comprise
CC contacting the cell with a labelled antibody that specifically recognises
CC and binds the protein product of at least one of the EGFR-RS, RYK,
CC TNFRSF25, TRPM7, KCP3 or KIAA 1883 genes, or with an immune effector cell
CC that specifically recognises and lyses a cell expressing at least one of
CC the EGFR-RS, RYK, TNFRSF25, TRPM7, UNC5H2, KCP3 and KIAA 1883 genes.
XX

SQ Sequence 1460 AA;
Query Match 100.0%; Score 45; DB 8; Length 1460;
Best Local Similarity 100.0%; Pred. No. 8.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
|||||||
Db 1142 RPQPPPP 1148

RESULT 42
ABR42259
ID ABR42259 standard; protein; 1734 AA.
XX
AC ABR42259;
XX
DT 28-JUL-2003 (first entry)
XX
DE Human MAST205 protein.
XX
KW Human; MAST205; protein kinase; enzyme; cytostatic; antidiabetic;
KW nootropic; neuroprotective; antiinflammatory; analgesic; gene therapy.
XX
OS Homo sapiens.
XX
PN WO2003033709-A2.
XX
PD 24-APR-2003.
XX
PF 16-OCT-2002; 2002WO-EP011551.
XX
PR 17-OCT-2001; 2001US-0329502P.
XX
PA (FARB) BAYER AG.
XX
PI Zhu Z;
XX
DR WPI; 2003-393534/37.
XX
PT New polynucleotide encoding a serine/threonine protein kinase
PT polypeptide, useful for preventing, ameliorating or treating diseases
PT associated with the protein kinase dysfunction, e.g. cancer, diabetes or
PT a CNS disorder.
XX
PS Disclosure; Fig 6; 149pp; English.
XX
CC The present sequence is that of human MAST205 protein. A novel human
CC serine/threonine protein kinase (see ABR42260) of the invention shows
CC homology to this sequence. The serine/threonine protein kinase can be
CC produced by recombinant methods and used to screen for reagents that
CC modulate its activity. Such reagents are used in the preparation of a
CC medicament for modulating the activity of the enzyme in a disease
CC selected from cancer, diabetes, central nervous system disorders and
CC chronic obstructive pulmonary disease (COPD) (claimed)

SQ Sequence 1734 AA;
Query Match 100.0%; Score 45; DB 6; Length 1734;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
|||||||
Db 9 RPQPPPP 15

RESULT 43
ABM78986
ID ABM78986 standard; protein; 1734 AA.
XX
AC ABM78986;
XX
DT 15-JAN-2004 (first entry)
XX
DE Modifier of beta-catenin (MBCAT) polypeptide GI 14149671.
XX
KW Modifier of beta-catenin; MBCAT; protein kinase; enzyme; cytostatic.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 512..785
FT /note= "kinase domain"
XX
PN WO2003074725-A2.
XX
PD 12-SEP-2003.
XX
PF 28-FEB-2003; 2003WO-US006294.
XX
PR 01-MAR-2002; 2002US-0361242P.
XX
PA (EXEL-) EXELIXIS INC.
XX
PI Costa MA, Gendreau SB, Dora EG, Nicoll M, Heuer TS;
XX
DR WPI; 2003-756771/71.
XX

PT Identifying a candidate beta-catenin pathway-modulating agent, for
PT diagnosing and treating cancer, comprises contacting an assay system
PT comprising a modifier of beta-catenin (MBCAT) polypeptide or nucleic acid
PT with a test agent.

PS Claim 13; Page 79-86; 114pp; English.

XX
CC The present sequence is that of human modifier of beta-catenin (MBCAT)
CC polypeptide GI 14149671. MBCATs are kinase proteins with protein kinase
CC domains. Genetic screens were designed to identify modifiers of the beta-
CC catenin pathway in Caenorhabditis elegans, and the C10C6.1 gene was
CC identified. Human orthologues, including the present sequence, were
CC subsequently obtained. MBCAT genes and polypeptides can be used to
CC identify MBCAT-modulating agents that are candidate therapeutic agents
CC for treatment of disorders associated with defective or impaired beta-
CC catenin and/or MBCAT function, such as an angiogenic, apoptotic or cell
CC proliferation disorder, e.g. breast, colon, head and neck, kidney, lung,
CC ovarian, prostate, skin or uterine cancer. MBCAT modulating agents
CC include antisense oligomers and RNAi that repress gene expression or
CC product activity. Transgenic non-human animals are useful as models of
CC disease and disorders implicating defective beta-catenin function

SQ Sequence 1734 AA;

Query Match 100.0%; Score 45; DB 7; Length 1734;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
|||||||

Db 9 RPQPPPP 15

RESULT 44

ADG20497

ID ADG20497 standard; protein; 1734 AA.

XX

AC ADG20497;

XX

DT 11-MAR-2004 (first entry)

XX

DE Human MAST205 protein sequence SEQ ID NO:30.

XX

KW diabetes; pre-diabetic; antidiabetic; insulin resistance; human; MAST205.

XX

OS Homo sapiens.

XX

PN WO2003103601-A2.

XX

PD 18-DEC-2003.

XX

PF 05-JUN-2003; 2003WO-US018046.

XX

PR 05-JUN-2002; 2002US-0386521P.

PR

PR 05-JUN-2002; 2002US-0386527P.

PR

PR 05-JUN-2002; 2002US-0386551P.

PR

PR 06-JUN-2002; 2002US-0386429P.

PR

PR 06-JUN-2002; 2002US-0386936P.

PR

PR 06-JUN-2002; 2002US-0386954P.

PR

PR 07-JUN-2002; 2002US-0387301P.

XX

PA (META-) METABOLEX INC.

XX

PI Allan B, Gregoire F, Lavan B, Moodie S, Waters S, Wong C;

XX

DR WPI; 2004-0622209/06.

DR N-PSDB; ADG20496.

XX

PT Identifying an agent for treating a diabetic or pre-diabetic individual, comprises selecting an agent that modulates the expression or activity of a diabetes-related polypeptide encoded by a diabetes-related polynucleotide.

XX

PS Claim 1; SEQ ID NO 30; 96pp; English.

XX

CC The present invention describes a method for identifying an agent for treating a diabetic or pre-diabetic individual. The method comprises contacting an agent to a mixture comprising a polypeptide encoded by a nucleic acid that hybridises under stringent conditions to a nucleic acid encoding any one of 10 sequences of 188-1734 amino acids (I) (see ADG20469, ADG20473, ADG20475, ADG20477, ADG20483, ADG20487, ADG20489, ADG20495, ADG20497 and ADG20501), and selecting an agent that modulates the expression or activity of the polypeptide or that binds to the polypeptide. Also described: (1) a method of treating a diabetic or pre-diabetic animal comprising administering to the animal a therapeutic amount of the agent identified by the method described above; (2) a method of introducing an expression cassette into a cell comprising introducing into the cell an expression cassette comprising a promoter operably linked to a polynucleotide encoding a polypeptide, where the polynucleotide hybridises under stringent conditions to a nucleic acid encoding (1); (3) a method of diagnosing an individual who has type 2 diabetes or is pre-diabetic comprising detecting in a sample from the individual the level of a polypeptide or a polynucleotide encoding the polypeptide, where the polynucleotide hybridises under stringent conditions to a nucleic acid encoding (I), and where a modulated level of the polypeptide or polynucleotide in the sample compared to a level of the polypeptide or polynucleotide in a lean individual or a previous sample from the individual indicates that the individual is diabetic or pre-diabetic; (4) an isolated nucleic acid encoding a polypeptide comprising a sequence of ADG20476 or ADG20494; (5) an expression cassette comprising the isolated nucleic acid; and (6) a host cell comprising the expression cassette. (I) has antidiabetic activity. The methods, polypeptides and polynucleotides are useful for identifying modulators of

CC the expression or activity of the polypeptides and polynucleotides. The CC modulators can be used for diagnosing and treating type 2 diabetes, pre-diabetes, or insulin resistance. The present sequence is used in the CC exemplification of the present invention.

XX

SQ Sequence 1734 AA;

Query Match 100.0%; Score 45; DB 8; Length 1734;

Best Local Similarity 100.0%; Pred. No. 1e+03;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7

Db 9 RPQPPPP 15

RESULT 45

ADQ15178

ID ADQ15178 standard; protein; 1734 AA.

XX

AC ADQ15178;

XX

DT 07-OCT-2004 (first entry)

XX

DE Human cancer related polypeptide, 21708.

XX

KW Cancer; angiogenic disorder; cytostatic; antiangiogenic; gene therapy; human.

XX

OS Homo sapiens.

XX

PN WO2004058153-A2.

XX

PD 15-JUL-2004.

XX

PF 16-DEC-2003; 2003WO-US040226.

XX

PR 20-DEC-2002; 2002US-0435108P.

PR

PR 23-DEC-2002; 2002US-0436443P.

PR

PR 07-JAN-2003; 2003US-0438498P.

PR

PR 31-JAN-2003; 2003US-0444370P.

PR

PR 06-FEB-2003; 2003US-0446031P.

PR

PR 11-MAR-2003; 2003US-0453635P.

PR

PR 25-MAR-2003; 2003US-0457199P.

PR

PR 10-APR-2003; 2003US-0462458P.

PR

PR 30-APR-2003; 2003US-0466732P.

PR

PR 08-MAY-2003; 2003US-0469184P.

PR

PR 19-MAY-2003; 2003US-0471663P.

PR

PR 03-JUN-2003; 2003US-0475472P.

PR

PR 12-JUN-2003; 2003US-0478150P.

PR

PR 23-JUN-2003; 2003US-0480631P.

PR

PR 15-JUL-2003; 2003US-0487369P.

PR

PR 29-JUL-2003; 2003US-0490866P.

PR

PR 02-SEP-2003; 2003US-0499614P.

PR

PR 09-OCT-2003; 2003US-0510081P.

PR

PR 06-NOV-2003; 2003US-0517742P.

XX

PA (MILL-) MILLENNIUM PHARM INC.

XX

PI Lightcap ES, Ecsedy JA, Hunter JJ, Macbeth KJ, Nestor MT;

XX

DR WPI; 2004-525772/50.

DR N-PSDB; ADQ15177.

XX

PT Identifying a compound capable of treating cancer or an angiogenic disorder by combining a compound to be tested with e.g., 15986, 2179 or 13249 polypeptide and detecting binding of the test compound to the polypeptide.

XX

PS Claim 1; SEQ ID NO 136; 576pp; English.

XX

CC The invention relates to a novel method for identifying a compound capable of treating cancer or an angiogenic disorder. The method

CC comprises: combining a compound to be tested with a cancer related
CC polypeptide; and detecting binding of the test compound to the
CC polypeptide. The invention further comprises: identifying a subject
CC having or at risk for developing cancer; and treating a subject having a
CC cancer or a cancer characterised by aberrant 15986, 2188, 20743, 9148,
CC 9151, 9791,44252, 14184, 42461, 8204, 7970, 25552, 21657, 26492, 2411,
CC 15088, 1905, 28899, 63380, 33935, 10480, 12686, 25501, 17694, 15701,
CC 53062, 49908, 21612, 38949, 6216, 46863, 9235, 2201, 6985, 9883, 12238,
CC 18057, 21617, 39228, 49928, 54476, 62113, 64316, 12264, 32362, 58198,
CC 2887, 3205, 8557, 9600, 9693, 44867, 53058, 55556, 57658, 2208, 10252,
CC 10302, 14218, 33877, 10317, 10485, 25964, 14815, 1363, 1397, 14827,
CC 21708, 3801, 64698, 2179 or 13249 cancer related polypeptide activity or
CC nucleic acid expression. The method and compounds have cytostatic and
CC antiangiogenic activities. The cancer related genes may be used in gene
CC therapy to treat disorders. The method is useful in identifying a
CC compound capable of treating cancer or an angiogenic disorder. This
CC sequence represents a human cancer related polypeptide used in the method
CC of the invention.

XX SQ Sequence 1734 AA;

Query Match 100.0%; Score 45; DB 8; Length 1734;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
| | | | |
Db 9 RPQPPPP 15

RESULT 46
AAM49177
ID AAM49177 standard; protein; 1798 AA.

XX AAM49177;

XX 14-JUN-2002 (first entry)

DE Human MAST205 (hMAST205).

XX Human; MAST205; PDZ domain; protocadherin LKC;
KW microtubule-associated spermatid manchette serine/threonine kinase 205;
KW cellular proliferation; cancer; tumour; drug screening; cytostatic;
KW anticancer; enzyme.

XX Homo sapiens.

XX Key Location/Qualifiers
FT Domain 1032. .1294
FT /label= PDZ domain
FT /note= "Involved in binding the C-terminal 4 residues of
FT protocadherin LKC (AAM49176, AAM49179)"

XX WO200198483-A1.

PN 27-DEC-2001.

XX 21-JUN-2001; 2001WO-JP005313.

XX 22-JUN-2000; 2000JP-00192693.

PR 09-FEB-2001; 2001JP-00034058.

XX (HELI-) HELIX RES INST.

XX Koga H, Takahashi N, Okazaki N, Masuho Y;

XX WPI; 2002-130792/17.

DR N-PSDB; ABL55198.

XX Anticancer function-carrying protocadherin-like protein, encoded gene and
PT their variants, applicable in cancer diagnosis e.g. as target for
PT evaluation of cancer cell proliferation and metastasis, and developing
PT drugs for cancer.

XX Claim 4; Page 81-94; 101pp; Japanese.
XX The invention relates to a human protocadherin-like protein designated
CC protocadherin LKC (AAM49176), to a human homologue of MAST205
CC (microtubule-associated spermatid manchette serine/threonine kinase 205)
CC designated hMAST205 (AAM49177) and to nucleic acids encoding them
CC (ABL55191, ABL55198). Protocadherin LKC and hMAST205 are able to bind to
CC each other via the interaction of the C-terminus of protocadherin LKC and
CC the PDZ domain of hMAST205, and are involved in the regulation of the
CC cellular proliferation. The protein and nucleic acid sequences of the
CC invention may be used in screening for drugs with anticancer activity.
CC Additionally, protocadherin LKC may be used in the diagnosis of cancer as
CC a target for the evaluation of cancer cell proliferation and metastasis,
CC where the risk of cancer is assessed when there is a reduction in
CC expression or a mutation in the protocadherin LKC gene. The present
CC sequence represents hMAST205

XX SQ Sequence 1798 AA;

Query Match 100.0%; Score 45; DB 5; Length 1798;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
| | | | |
Db 9 RPQPPPP 15

RESULT 47
AAE16277
ID AAE16277 standard; protein; 1798 AA.

XX AAE16277;

XX 26-MAR-2002 (first entry)

XX Human kinase PKIN-23 protein.

XX Human; kinase; PKIN-23; cancer; leukaemia; adenocarcinoma; osteoporosis;
KW immune disorder; atherosclerosis; Crohn's disease; Hodgkin's disease;
KW Acquired Immune Deficiency Syndrome; AIDS; Addison's disease; anaemia;
KW allergy; asthma; adult respiratory distress syndrome; multiple sclerosis;
KW autoimmune thyroiditis; bronchitis; diabetes mellitus; osteoarthritis;
KW Good pasture's syndrome; Graves' disease; pancreatitis; psoriasis;
KW rheumatoid arthritis; ulcerative colitis; cirrhosis; Cushing's syndrome;
KW hepatitis; hypothyroidism; cerebral palsy; cataract; angina pectoris;
KW cardiovascular disease; hypertension; vasculitis; myocarditis; obesity;
KW congestive heart failure; ischaemic heart disease; lung tumour; gout;
KW fatty liver; Niemann-Pick's disease; gene therapy; chromosome 13.

XX Homo sapiens.

XX Key Location/Qualifiers
FH Domain 512. .785
FT /note= "Eukaryotic protein kinase domain"

FT Domain 1104. .1191
FT /label= PDZ_domain

XX WO200196547-A2.

PN 20-DEC-2001.

XX 14-JUN-2001; 2001WO-US019444.

XX 15-JUN-2000; 2000US-0212073P.

PR 23-JUN-2000; 2000US-0213467P.

PR 30-JUN-2000; 2000US-0215651P.

PR 07-JUL-2000; 2000US-0216605P.

PR 13-JUL-2000; 2000US-0218372P.

PR 25-AUG-2000; 2000US-0228056P.

XX (INCY-) INCYTE GENOMICS INC.

PA

XX Yue H, Lal P, Bandman O, Borowsky ML, Au-Young J, Lu Y;
PI Gandhi AR, Tribouley CM, Walia NK, Yao MG, Lu DAM, Greenwald SR;
PI Ramkumar J, Griffin JA, Kearney L, Burford N, Nguyen DB, Tang YT;
PI Baughn MR, He A, Thornton M, Hafalia A, Patterson C, Gururajan R;
PI Lo TP, Khan F, Recipon SA, Azimzai Y, Policky JL, Ding L;
PI Grether M, Elliott VS, Thangavelu K, Batra S, Ison CH;
XX WPI; 2002-090207/12.
DR N-PSDB; AAD26470.
DR
XX
PT New polypeptides, useful for diagnosing, treating or preventing disorders
PT of growth and development, cardiovascular and lipid, and diseases such as
PT cancer, comprise human kinase polypeptides.
XX
PS Claim 1; Page 169-172; 197pp; English.
XX
CC The invention relates to human kinase PKIN proteins and their
CC corresponding cDNAs. A composition containing PKIN agonist is useful for
CC treating a disease or condition associated with decreased expression of
CC PKIN and a composition comprising PKIN antagonist is useful for treating
CC a disease or condition associated with overexpression of PKIN. The
CC disorders include cancer (leukaemia, adenocarcinoma, lymphoma, melanoma,
CC myeloma, sarcoma, teratocarcinoma, Hodgkin's disease); immune disorder
CC (Acquired Immune Deficiency Syndrome (AIDS), asthma, Addison's disease,
CC atherosclerosis, anaemia, allergies, adult respiratory distress syndrome,
CC autoimmune thyroiditis, gout, bronchitis, Crohn's disease, diabetes
CC mellitus, multiple sclerosis, Good pasture's syndrome, Graves' disease,
CC osteoarthritis, osteoporosis, pancreatitis, psoriasis, Reiter's syndrome,
CC rheumatoid arthritis, Sjogren's syndrome, uveitis, ulcerative colitis,
CC bacterial, parasitic, fungal, viral, protozoal and helminthic infections)
CC growth and development disorders (arteriosclerosis, cirrhosis, hepatitis,
CC Cushing's syndrome, hypothyroidism, cerebral palsy, cataracts); cardio
CC vascular disease (arteriovenous fistula, hypertension, vasculitis,
CC aneurysms, congestive heart failure, angina pectoris, myocarditis,
CC ischaemic heart disease, chronic bronchitis, lung tumours); lipid
CC disorder (fatty liver, Fabry's disease, Niemann-Pick's disease,
CC hypocholesterolaemia, obesity). PKIN DNA is useful for assessing toxicity
CC of a test compound and in gene therapy. The present sequence is human
CC PKIN-23 protein. Human PKIN-23 gene is located on chromosome 13
XX
SQ Sequence 1798 AA;

Query Match 100.0%; Score 45; DB 5; Length 1798;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
Db 9 RPQPPPP 15

RESULT 48
ADJ96613
ID ADJ96613 standard; protein; 1798 AA.
XX
AC ADJ96613;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human microtubule associated ser/thr protein kinase MAST205 SeqID 70.
XX
KW kinase; human; SNP; single nucleotide polymorphism;
KW tyrosine protein kinase; serine/threonine protein kinase; PTK; STK;
KW gene therapy; cancer; immune-related disease; cardiovascular disease;
KW brain; neuronal associated disease; metabolic; inflammatory disorder;
KW cytosstatic; neuroprotective; immunomodulator; antiinflammatory; enzyme;
KW microtubule associated ser/thr protein kinase; MAST205.
XX
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT Misc-difference 9

FT
FT /note= "Wild type Arg substituted for STOP by single
FT nucleotide polymorphism"
FT Misc-difference 1551
FT /note= "Wild type Asp substituted for Gly by single
FT nucleotide polymorphism"
FT Misc-difference 1768
FT /note= "Wild type Ser substituted for Phe by single
FT nucleotide polymorphism"
FT
XX
PN WO2004006838-A2.
XX
XX
PD 22-JAN-2004.
XX
PF 15-JUL-2003; 2003WO-US021730.
XX
PR 15-JUL-2002; 2002US-0395632P.
XX
PA (SUGE-) SUGEN INC.
XX
PI Whyte D, Manning G, Caenepeel S;
XX
DR WPI; 2004-122753/12.
DR N-PSDB; ADJ96547.
XX
PT New nucleic acid molecule encoding a kinase polypeptide, useful for
PT preparing a composition for treating diseases or disorders, e.g., cancer,
PT or neurological, immunological or inflammatory disorders.
XX
PS Claim 1; SEQ ID NO 70; 366pp; English.
XX
CC This invention relates to a novel isolated, enriched or purified nucleic
CC acid molecule that encodes a kinase polypeptide. Specifically, it relates
CC to human tyrosine and serine/threonine protein kinases (PRK's and STK's),
CC as well as protein kinase-like enzymes. The present invention describes
CC screening methods to identify agonists, antagonists and antibodies that
CC can be used to modulate the activity or function of the mammalian kinase
CC enzymes. As such, these compositions can be used for gene therapy
CC purposes to treat diseases or disorders including cancer, immune-related
CC diseases, cardiovascular disease, brain or neuronal associated disease,
CC metabolic and inflammatory disorders. Accordingly, they exhibit
CC cytosstatic, neuroprotective, immunomodulator and antiinflammatory
CC activities. This polypeptide sequence is a human kinase protein sequence
CC of the invention.
XX
SQ Sequence 1798 AA;

Query Match 100.0%; Score 45; DB 8; Length 1798;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
Db 9 RPQPPPP 15

RESULT 49
ADX91705
ID ADX91705 standard; protein; 1805 AA.
XX
AC ADX91705;
XX
DT 21-APR-2005 (first entry)
XX
DE Plant full length insert polypeptide seqid 54369.
XX
KW plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content.

XX OS Unidentified.
XX US2004034888-A1.
PN 19-FEB-2004.
XX 28-APR-2003; 2003US-00425114.
PF 06-MAY-1999; 99US-00304517.
XX 05-NOV-2001; 2001US-00985678.
XX (LIUJ/) LIU J.
PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA (TABA/) TABASKA J E.
PA (CAOY/) CAO Y.
XX
PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JB, Cao Y;
XX WPI; 2004-180133/17.
DR
XX New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX Claim 1; SEQ ID NO 54369; 15pp; English.
XX
CC The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This is the amino acid sequence of a plant full length insert
CC polypeptide that can be used in the recombinant DNA construct of the
CC invention.
XX
SQ Sequence 1805 AA;
Query Match 100.0%; Score 45; DB 8; Length 1805;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 RPQPPPP 7
Db 17 RPQPPPP 23
RESULT 50
ADH47759
ID ADH47759 standard; protein; 4952 AA.
XX
AC ADH47759;
XX
DT 25-MAR-2004 (first entry)
XX
DE NOV18 protein, SEQ ID 56.
XX
KW Antidiabetic; anorectic; cardiant; hypotensive; antiarteriosclerotic;
KW anorectic; virucide; antibacterial; fungicide; protozoacide; nootropic;
KW neuroprotective; antiparkinsonian; anticonvulsant; osteopathic;

KW antiarthritic; antiinflammatory; dermatological; antiasthmatic;
KW antilipaemic; Gene therapy; human; metabolic disorder; diabetes; obesity;
KW viral infection; bacterial infection; fungal infection;
KW helminthic infection; protozoal infection; anorexia; cancer;
KW cardiovascular disease; neurodegenerative disorder; Alzheimer's disease;
KW Parkinson's disease; epilepsy; immune disorder; haematopoietic disorder;
KW inflammatory skin disorder; asthma; dyslipidaemia; NOV18;
KW ARL-like protein; chromosome 12q12-q14.
XX
OS Homo sapiens.
XX WO200268647-A2.
XX 06-SEP-2002.
XX 16-JAN-2002; 2002WO-US001311.
PR 16-JAN-2001; 2001US-0261376P.
PR 18-JAN-2001; 2001US-0262454P.
PR 18-JAN-2001; 2001US-0262587P.
PR 31-JAN-2001; 2001US-0265530P.
PR 14-FEB-2001; 2001US-0268595P.
PR 28-FEB-2001; 2001US-0272409P.
PR 16-MAR-2001; 2001US-0276777P.
PR 17-MAY-2001; 2001US-0291672P.
PR 27-SEP-2001; 2001US-0325306P.
PR 18-OCT-2001; 2001US-0330336P.
PR 09-NOV-2001; 2001US-0345202P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Padigaru M, Alsobrook JP, Colman SD, Spytek KA, Boldog F;
PI Vernet CAM, Li L, Shenoy S, Casman S, Guo X, Edinger S;
PI Macdougall J, Malyankar U, Patturajan M, Shimkets RA, Pena C;
PI Tchernev V, Zerhusen BD, Millett I, Miller C, Lepley DM, Smithson G;
PI Baumgartner J, Herrmann J, Peyman JA, Gorman L, Mezes P, Kekuda R;
PI Taupier RJ, Gerlach V, Grosse WM, Liu X, Ellerman K, Rothenberg M;
PI Stone DJ, Burgess CE;
XX WPI; 2002-698671/75.
DR N-PSDB; ADH47758.
XX
PT New isolated NOVX polypeptides and polynucleotides, useful for
PT preventing, diagnosing or treating NOVX-associated disorders e.g.
PT osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
PT asthma, or infections.
XX
PS Claim 1; Page 159-160; 380pp; English.
XX
CC The present invention relates to novel proteins (I) referred to as NOVX,
CC where X is any number from 1 to 18, and their coding sequences (II). The
CC proteins and their coding sequences are useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease,
CC preferably a NOVX-associated disorder such as metabolic disorders,
CC diabetes, obesity, infectious diseases (viral, bacterial, fungal,
CC helminthic, and protozoal), anorexia, cancer, cardiovascular diseases
CC (hypertension, atherosclerosis), neurodegenerative disorders, Alzheimer's
CC disease, Parkinson's disease, epilepsy, immune disorders
CC (osteoarthritis), haematopoietic disorders, inflammatory skin disorders,
CC asthma, and various dyslipidaemias. NOV18 comprises a ARL-like protein
CC and maps to chromosome 12q12-q14.
XX
SQ Sequence 4952 AA;
Query Match 100.0%; Score 45; DB 5; Length 4952;
Best Local Similarity 100.0%; Pred. No. 2.5e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 RPQPPPP 7
Db 1806 RPQPPPP 1812

Search completed: April 6, 2006, 09:32:23
Job time : 109.289 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2006, 09:35:13 ; Search time 12.7105 Seconds
(without alignments)
52.989 Million cell updates/sec

Title: US-10-632-388-305

Perfect score: 45

Sequence: 1 RPPPPPP 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : PIR 80:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	125	A53692	synapsin I - mouse
2	45	100.0	691	A25704	synapsin I - rat
3	45	100.0	704	A30411	synapsin Ia - rat
4	45	100.0	705	A35363	synapsin I splice
5	45	100.0	706	E30411	synapsin Ia - bovi
6	45	100.0	4957	T03455	ALR protein - huma
7	45	100.0	5262	T03454	ALR protein - huma
8	42	93.3	139	T05847	hypothetical prote
9	42	93.3	320	AE2842	conserved hypothet
10	42	93.3	359	F97619	hypothetical prote
11	42	93.3	645	A71416	hypothetical prote
12	42	93.3	736	I51691	dishevelled homolo
13	42	93.3	1278	T30188	Niemann-Pick C dis
14	42	93.3	1745	A46431	tight junction-ass
15	41	91.1	159	A83295	hypothetical prote
16	41	91.1	248	JQ1682	infected cell prot
17	41	91.1	320	A39724	homeotic protein H
18	41	91.1	324	A41786	mRNA-binding prote
19	41	91.1	634	T00388	hypothetical prote
20	41	91.1	780	T00366	hypothetical prote
21	41	91.1	972	T47922	anthranilate phosp
22	41	91.1	979	A35913	regulatory factor
23	41	91.1	1029	T30351	mucin-like protein
24	41	91.1	2907	A57278	fibrillin-2 precur
25	41	91.1	2918	A54105	fibrillin-2 precur
26	41	91.1	3144	A46068	Huntington disease
27	40	88.9	70	A96650	protein F24O1.6 li
28	40	88.9	196	B48232	cysteine-rich exte
29	40	88.9	209	A48232	cysteine-rich exte

30	40	88.9	215	2	S34163	homeotic protein H
31	40	88.9	221	2	T07176	extensin homolog -
32	40	88.9	242	2	S54156	extensin-like prot
33	40	88.9	302	2	C86480	33.2K hypothetical
34	40	88.9	303	2	JQ1386	hypothetical 33K p
35	40	88.9	312	2	B29350	env polypeptin, r
36	40	88.9	334	2	E70519	probable oxidoredu
37	40	88.9	334	2	T33836	hypothetical prote
38	40	88.9	346	2	JC5833	transcription fact
39	40	88.9	349	2	T49791	hypothetical prote
40	40	88.9	356	1	VCVWB6	env polypeptin pr
41	40	88.9	356	2	S31428	envelope protein -
42	40	88.9	408	1	VCMVSR	env polypeptin -
43	40	88.9	408	2	S50899	betaB inhibin prec
44	40	88.9	409	2	A21170	gp55 protein - Fri
45	40	88.9	411	2	S34164	homeotic protein H
46	40	88.9	413	2	T04520	hypothetical prote
47	40	88.9	416	2	S27198	homeotic protein H
48	40	88.9	416	2	T34279	hypothetical prote
49	40	88.9	417	2	S47539	homeotic protein H
50	40	88.9	422	2	A83184	probable protein m
51	40	88.9	456	2	S23104	choline kinase - h
52	40	88.9	469	2	I37451	HBF-G2 (HFK-2) pro
53	40	88.9	476	2	A54743	transcription fact
54	40	88.9	480	2	JH0672	brain factor 1 pro
55	40	88.9	481	2	T27975	hypothetical prote
56	40	88.9	489	2	T26069	hypothetical prote
57	40	88.9	495	1	S31223	transcription fact
58	40	88.9	533	1	S56552	calcium-dependent
59	40	88.9	546	2	T02029	DNA-binding protei
60	40	88.9	549	2	C87719	protein R119.6 [im
61	40	88.9	588	2	T45564	hypothetical prote
62	40	88.9	610	2	S35049	mucin JER57 - huma
63	40	88.9	610	2	S55643	glycoprotein L - e
64	40	88.9	617	2	I56530	gene VGF protein -
65	40	88.9	636	1	VCVWFS	env polypeptin -
66	40	88.9	640	1	VCVWM1	env polypeptin -
67	40	88.9	640	1	VCMVRV	env polypeptin pr
68	40	88.9	655	2	AD2422	DNA polymerase III
69	40	88.9	655	2	S40521	FKHR protein - hum
70	40	88.9	656	2	B49423	semaphorin I - fru
71	40	88.9	669	2	T38029	matrix metalloprot
72	40	88.9	756	2	T27642	hypothetical prote
73	40	88.9	785	2	S75819	general secretion
74	40	88.9	839	2	T04859	extensin homolog F
75	40	88.9	963	2	A55926	DNA binding protei
76	40	88.9	1026	1	A40315	maternal effect pr
77	40	88.9	1174	2	T43051	protein kinase C (
78	40	88.9	1201	2	A57369	anillin - fruit fl
79	40	88.9	1213	2	S16356	ovo protein - frui
80	40	88.9	1234	2	T30160	hypothetical prote
81	40	88.9	1240	2	JC5209	insulin receptor s
82	40	88.9	1242	2	JS0670	insulin receptor s
83	40	88.9	1316	2	T00381	KIAA0633 protein -
84	40	88.9	1403	2	S24548	homeotic protein p
85	40	88.9	1460	1	EDBEIF	immediate-early pr
86	40	88.9	1541	2	T02831	AAA protein L4171.
87	40	88.9	1840	2	T30250	GT1 protein - mous
88	40	88.9	2649	2	T51023	hypothetical prote
89	39	86.7	126	2	H86347	hypothetical prote
90	39	86.7	147	2	S37485	gene msg1 protein
91	39	86.7	148	2	S39206	proline-rich prote
92	39	86.7	148	2	E75300	conserved hypothet
93	39	86.7	161	2	T04981	hypothetical prote
94	39	86.7	161	2	H85120	probable proline-r
95	39	86.7	198	2	T35752	hypothetical prote
96	39	86.7	199	2	S14981	extensin class I (
97	39	86.7	213	2	C86575	hypothetical prote
98	39	86.7	213	2	B72049	hypothetical prote
99	39	86.7	219	2	T10563	proline-rich prote
100	39	86.7	219	2	T27105	hypothetical prote
101	39	86.7	248	2	F91197	EspF protein [impo
102	39	86.7	248	2	B86044	espF [imported] -

103	39	86.7	278	2	A49266	fas ligand - rat	176	38	84.4	633	2	S62057	proline-rich prote
104	39	86.7	279	2	A53062	Fas ligand - mouse	177	38	84.4	982	1	GNLJH2	pol polyprotein -
105	39	86.7	281	2	I38707	Fas ligand - human	178	38	84.4	1091	2	S33596	protein-tyrosine k
106	39	86.7	306	2	T06435	ribonuclease S5 ho	179	38	84.4	1122	2	T47424	hypothetical prote
107	39	86.7	333	2	C48423	homeotic protein e	180	38	84.4	1338	2	T30565	MAP kinase kinase
108	39	86.7	348	2	T47494	hypothetical prote	181	38	84.4	1474	2	T18281	hypothetical prote
109	39	86.7	373	2	T51916	hypothetical prote	182	38	84.4	1776	2	G86280	protein TSG21.13 {
110	39	86.7	382	2	B86575	hypothetical prote	183	38	84.4	1872	2	T00339	hypothetical prote
111	39	86.7	382	2	A72049	hypothetical prote	184	38	84.4	2584	2	T24158	hypothetical prote
112	39	86.7	383	2	AE2295	hypothetical prote	185	38	84.4	2606	2	T24157	hypothetical prote
113	39	86.7	393	2	T32533	hypothetical prote	186	37	82.2	75	2	S14973	extensin class I (
114	39	86.7	429	1	FOLJCN	gag polyprotein -	187	37	82.2	104	2	H96000	hypothetical prote
115	39	86.7	429	1	FOLJGH	probable gag poly	188	37	82.2	107	2	E72710	hypothetical prote
116	39	86.7	429	2	S06073	gag polyprotein -	189	37	82.2	142	2	A11967	hypothetical prote
117	39	86.7	439	2	S51939	chitinase (EC 3.2.	190	37	82.2	143	2	D83148	hypothetical prote
118	39	86.7	464	2	S22697	extensin - Volvox	191	37	82.2	145	2	T22693	hypothetical prote
119	39	86.7	508	2	T45867	hypothetical prote	192	37	82.2	160	2	D48232	cysteine-rich exte
120	39	86.7	516	2	T00974	probable SF16 prot	193	37	82.2	161	2	B48232	cysteine-rich exte
121	39	86.7	574	2	T43556	Wiskott-Aldrich sy	194	37	82.2	161	2	F72593	hypothetical prote
122	39	86.7	574	2	T38819	wiskott-aldrich sy	195	37	82.2	165	2	C48232	cysteine-rich exte
123	39	86.7	610	2	T06690	galactonolactone d	196	37	82.2	168	2	T48343	hypothetical prote
124	39	86.7	643	2	T27429	hypothetical prote	197	37	82.2	168	2	T48375	transcription co-a
125	39	86.7	653	2	T08984	auxin response fac	198	37	82.2	215	2	B75281	peptidyl-trNA hydr
126	39	86.7	715	1	TNBE77	77K alpha trans-in	199	37	82.2	235	2	T31424	C-terminal domain-
127	39	86.7	742	2	F84643	hypothetical prote	200	37	82.2	266	1	C1PGL	calpain (EC 3.4.22
128	39	86.7	760	2	T06291	extensin homolog T	201	37	82.2	266	1	C1RBL	calpain (EC 3.4.22
129	39	86.7	786	2	T49414	related to ahmpl p	202	37	82.2	268	1	C1HUL	calpain (EC 3.4.22
130	39	86.7	788	2	T25061	hypothetical prote	203	37	82.2	271	2	A70415	molybdopterin bios
131	39	86.7	814	2	G02390	disintegrin-like m	204	37	82.2	294	2	T34537	hypothetical prote
132	39	86.7	868	2	T20239	hypothetical prote	205	37	82.2	329	2	T01333	hypothetical prote
133	39	86.7	872	2	T50369	probable serine/th	206	37	82.2	354	2	A48931	transmembrane glyc
134	39	86.7	894	2	F84870	hypothetical prote	207	37	82.2	361	2	S19552	potassium channel
135	39	86.7	969	2	T15446	hypothetical prote	208	37	82.2	367	2	T01751	gibberellin 20-oxi
136	39	86.7	992	1	GNWVR3	structural polypro	209	37	82.2	376	1	QQBEW2	UL53 protein - hum
137	39	86.7	1071	2	S48378	probable membrane	210	37	82.2	387	2	T21370	hypothetical prote
138	39	86.7	1111	2	T23047	hypothetical prote	211	37	82.2	390	2	AC2169	molybdopterin bios
139	39	86.7	1147	2	T42627	ADP-ribosylation f	212	37	82.2	391	2	T52532	molybdopterin synt
140	39	86.7	1201	2	G86441	unknown protein [i	213	37	82.2	392	2	S76179	hypothetical prote
141	39	86.7	1201	2	T29329	hypothetical prote	214	37	82.2	392	2	G70594	probable MOEB- hom
142	39	86.7	1247	2	T45743	hypothetical prote	215	37	82.2	396	1	S17929	transcription init
143	39	86.7	1307	2	S53412	probable membrane	216	37	82.2	400	2	S58222	PQ-rich protein -
144	39	86.7	1443	2	S05979	steroid hormone re	217	37	82.2	425	2	S33045	hypothetical prote
145	39	86.7	1575	2	S68448	synaptojanin, 170K	218	37	82.2	440	1	XXHUN	phosphatidylcholin
146	39	86.7	1584	2	T18276	protein-tyrosine k	219	37	82.2	440	2	JC1502	phosphatidylcholin
147	39	86.7	1599	2	T15854	hypothetical prote	220	37	82.2	473	2	S50755	hypothetical prote
148	39	86.7	1997	1	S12050	protein-tyrosine-p	221	37	82.2	481	2	E83062	deoxyribodipyrimid
149	39	86.7	3051	2	S42373	hypothetical prote	222	37	82.2	487	2	B83120	probable outer mem
150	39	86.7	7962	2	I38346	elastic titin - hu	223	37	82.2	533	2	B56110	tyrosine phosphor
151	38	84.4	106	2	A30996	orf-X protein - si	224	37	82.2	621	2	D96554	hypothetical prote
152	38	84.4	112	1	ASLJX3	vpu protein - simi	225	37	82.2	646	2	S55048	protein-tyrosine k
153	38	84.4	112	1	ASLJXC	vpu protein - huma	226	37	82.2	655	2	A86153	hypothetical prote
154	38	84.4	112	1	ASLJST	vpu protein - simi	227	37	82.2	679	2	A42073	potassium channel
155	38	84.4	112	1	ASLJX2	vpu protein - huma	228	37	82.2	681	2	F85062	hypothetical prote
156	38	84.4	112	2	S12155	vpv protein - huma	229	37	82.2	695	2	A87494	topoisomerase IV,
157	38	84.4	112	2	S53094	vpv protein - huma	230	37	82.2	751	2	T15403	hypothetical prote
158	38	84.4	112	2	S03066	gene X protein - h	231	37	82.2	760	2	F86387	probable Pto kinas
159	38	84.4	152	2	T00877	hypothetical prote	232	37	82.2	769	2	I56546	Shaw type potassiu
160	38	84.4	156	1	GNVQL2	genome-linked prot	233	37	82.2	793	1	KXRTF	urin (EC 3.4.21.7
161	38	84.4	156	1	GNVQLL	genome-linked prot	234	37	82.2	798	2	T21369	hypothetical prote
162	38	84.4	156	1	GNVQWA	genome-linked prot	235	37	82.2	806	1	JN0612	protein kinase B-r
163	38	84.4	219	2	G82611	hypothetical prote	236	37	82.2	807	1	I51153	protein kinase B-r
164	38	84.4	280	2	T50583	ABC-type transport	237	37	82.2	815	2	T15402	hypothetical prote
165	38	84.4	304	1	A60472	uracil-DNA glycosy	238	37	82.2	897	2	A54696	EGF receptor subst
166	38	84.4	425	2	T16433	hypothetical prote	239	37	82.2	908	2	T25035	hypothetical prote
167	38	84.4	450	2	A53771	transcription fact	240	37	82.2	1006	2	G86292	hypothetical prote
168	38	84.4	457	2	JC6551	chitinase (EC 3.2.	241	37	82.2	1160	2	F88369	protein unc-52 lim
169	38	84.4	501	2	A42030	alpha-globin trans	242	37	82.2	1255	1	A24571	hypothetical prote
170	38	84.4	502	2	C56205	transcription fact	243	37	82.2	1313	2	F96673	hypothetical prote
171	38	84.4	502	2	B42030	alpha-globin trans	244	37	82.2	2115	2	S38480	nonstructural prot
172	38	84.4	542	2	A82965	hypothetical prote	245	37	82.2	2295	2	C88369	protein unc-52 lim
173	38	84.4	547	2	C82655	hypothetical prote	246	37	82.2	2783	1	A41948	alpha-fetoprotein
174	38	84.4	618	2	T05518	hypothetical prote	247	37	82.2	3078	2	T28432	variant-specific s
175	38	84.4	626	1	NDBE61	exonuclease (EC 3.	248	37	82.2	3375	2	T19821	hypothetical prote

249 37 82.2 26926 1 I38344 titin, cardiac mus
250 36 80.0 75 2 T26916 hypothetical prote
251 36 80.0 80 2 PQ0477 pistil extensin-li
252 36 80.0 98 2 T42171 hypothetical prote
253 36 80.0 117 2 D40750 proline-rich prote
254 36 80.0 122 2 PQ0452 extensin-like prot
255 36 80.0 151 2 JQ1686 extensin-like prote
256 36 80.0 160 2 C72501 hypothetical prote
257 36 80.0 191 2 JQ1919 hypothetical 21.3K
258 36 80.0 211 2 S28304 hypothetical prote
259 36 80.0 212 2 B36298 proline-rich prote
260 36 80.0 229 1 W4WL8 E4 protein - human
261 36 80.0 231 2 B83031 conserved hypothet
262 36 80.0 232 2 T10646 hypothetical prote
263 36 80.0 235 1 WZBE23 gene 23 protein -
264 36 80.0 239 2 T16159 hypothetical prote
265 36 80.0 240 2 F87537 TonB protein, prob
266 36 80.0 245 1 W4WL5 E4 protein - human
267 36 80.0 245 1 W4WLB5 E4 protein - human
268 36 80.0 249 2 T04939 hypothetical prote
269 36 80.0 278 2 B39066 proline-rich prote
270 36 80.0 281 2 S65692 E4 protein - human
271 36 80.0 283 2 S42634 homeoprotein, IDX-
272 36 80.0 296 2 H87511 hypothetical prote
273 36 80.0 299 2 G70784 probable mmps3 pro
274 36 80.0 304 1 W4WL47 E4 protein - human
275 36 80.0 309 2 S10889 proline-rich prote
276 36 80.0 309 2 B87576 hypothetical prote
277 36 80.0 312 2 A56911 TRADD protein - hu
278 36 80.0 336 2 S29507 metavinculin - pig
279 36 80.0 340 2 T19105 phosphate carrier
280 36 80.0 342 2 T26677 hypothetical prote
281 36 80.0 351 2 B34768 ORF5 protein - Orf
282 36 80.0 356 2 A34350 phosphate transpor
283 36 80.0 361 2 B53737 phosphate carrier
284 36 80.0 361 2 D53737 phosphate carrier
285 36 80.0 362 2 A53737 phosphate carrier
286 36 80.0 362 2 C53737 phosphate carrier
287 36 80.0 363 2 B84565 probable spliceoso
288 36 80.0 363 2 T16755 hypothetical prote
289 36 80.0 367 2 I39172 cyclin-dependent k
290 36 80.0 379 2 A47659 farnesyl-protein t
291 36 80.0 379 2 T19069 hypothetical prote
292 36 80.0 387 2 F96798 hypothetical prote
293 36 80.0 388 2 JC5437 spliceosome-associ
294 36 80.0 392 1 PIHUB6 salivary proline-r
295 36 80.0 392 2 T15755 hypothetical prote
296 36 80.0 392 2 A86206 hypothetical prote
297 36 80.0 393 2 PQ0479 pistil extensin-li
298 36 80.0 397 2 S09813 hypothetical prote
299 36 80.0 403 2 S52796 prpL2 protein - hu
300 36 80.0 409 2 T24543 hypothetical prote
301 36 80.0 426 2 JQ1696 pistil extensin-li
302 36 80.0 432 2 AB2222 twitching motility
303 36 80.0 444 2 E83802 hypothetical prote
304 36 80.0 464 2 C84686 probable AP2 domai
305 36 80.0 465 2 A44498 radial spoke prote
306 36 80.0 468 1 FOMVMU gag polypotein -
307 36 80.0 538 1 FOMVIM gag polypotein -
308 36 80.0 538 1 FOMVM gag polypotein -
309 36 80.0 547 2 T02274 hypothetical prote
310 36 80.0 549 2 I37386 fas-activated seri
311 36 80.0 557 2 S23429 protein kinase ERK
312 36 80.0 560 2 I59302 brain specific Na+
313 36 80.0 564 2 T45866 hypothetical prote
314 36 80.0 602 2 T45760 hypothetical prote
315 36 80.0 606 1 UZAD12 terminal protein p
316 36 80.0 618 2 JC4366 transcription fact
317 36 80.0 651 2 T21175 hypothetical prote
318 36 80.0 671 2 T02504 hypothetical prote
319 36 80.0 721 2 C84677 probable membrane
320 36 80.0 778 2 T30430 hypothetical prote
321 36 80.0 782 2 T48722 hypothetical prote

322 36 80.0 816 2 F31277 regulatory protein
323 36 80.0 828 2 T06133 hypothetical prote
324 36 80.0 846 1 QOBEC3 HQRf1 protein - hu
325 36 80.0 882 2 S41034 hypothetical prote
326 36 80.0 893 2 G88551 protein M01A8.2 [i
327 36 80.0 903 2 T09143 alpha-glucosidase
328 36 80.0 913 2 JC5463 alpha-glucosidase
329 36 80.0 981 1 FOMVGM gag-abl polyprotei
330 36 80.0 1000 2 S44898 ZK1236.3 protein -
331 36 80.0 1003 2 T34066 hypothetical prote
332 36 80.0 1008 2 T04462 hypothetical prote
333 36 80.0 1010 2 A33509 vinculin - Caenorh
334 36 80.0 1013 2 T33470 hypothetical prote
335 36 80.0 1066 2 T10108 vinculin - mouse
336 80.0 1069 2 D85383 hypothetical prote
337 36 80.0 1072 2 T50949 verprolin related
338 36 80.0 1100 2 T21544 hypothetical prote
339 36 80.0 1123 2 A39962 kinase-related tra
340 36 80.0 1130 1 TVHUA protein-tyrosine k
341 36 80.0 1134 1 A35955 meta-vinculin - hu
342 36 80.0 1135 1 A29997 meta-vinculin - ch
343 36 80.0 1151 2 T18535 high molecular mas
344 36 80.0 1166 2 T13958 synGAP-b1 protein
345 36 80.0 1219 2 S54570 probable membrane
346 36 80.0 1245 1 VHWVB structural polypro
347 36 80.0 1245 1 VHWVB2 structural polypro
348 36 80.0 1249 2 A56511 myosin I myoA - Em
349 36 80.0 1249 2 T14270 Ras-GTPase activat
350 36 80.0 1260 2 T01334 hypothetical prote
351 36 80.0 1262 2 T13353 protein stn-B - fr
352 36 80.0 1293 2 T14259 ras GTPase-activat
353 36 80.0 1367 2 T13703 tamA protein - fru
354 36 80.0 1505 2 S28079 DNA-directed DNA p
355 36 80.0 1596 2 A35927 190K DNA-binding p
356 36 80.0 1733 1 B45344 probable nuclear a
357 36 80.0 1801 2 T26774 hypothetical prote
358 36 80.0 1958 2 B40505 hypothetical prote
359 36 80.0 2148 1 A56081 insulin receptor -
360 36 80.0 2182 2 T28634 variant-specific s
361 36 80.0 2715 2 T13049 eyelid - fruit fly
362 36 80.0 3005 2 T08841 polypotein - dour
363 35 77.8 67 2 S14971 extensin class I (
364 35 77.8 76 2 E97763 protein transport
365 35 77.8 82 2 T30135 hypothetical prote
366 35 77.8 90 2 S14972 extensin class I (
367 35 77.8 92 2 B48831 vitelline membrane
368 35 77.8 100 2 T17126 hypothetical prote
369 35 77.8 107 2 S20693 hypothetical prote
370 35 77.8 107 2 C87312 hypothetical prote
371 35 77.8 108 2 T05231 conserved hypothet
372 35 77.8 109 2 AG2992 hypothetical prote
373 35 77.8 109 2 C98291 hypothetical prote
374 35 77.8 131 2 B96566 extensin class I (
375 35 77.8 132 2 S14970 hypothetical prote
376 35 77.8 132 2 H96708 proline-rich prote
377 35 77.8 134 2 JC5572 hypothetical prote
378 35 77.8 138 2 C96734 hypothetical prote
379 35 77.8 139 2 T46186 collagen-related p
380 35 77.8 142 2 B41132 hypothetical prote
381 35 77.8 144 2 E86364 hypothetical prote
382 35 77.8 148 2 D72652 hypothetical prote
383 35 77.8 148 2 AD2681 transcription regu
384 35 77.8 148 2 B97463 probable repressor
385 35 77.8 149 2 A41132 collagen-related p
386 35 77.8 154 2 PQ0476 pistil extensin-li
387 35 77.8 156 2 C84556 probable bZIP tran
388 35 77.8 159 2 A72680 hypothetical prote
389 35 77.8 163 2 A29356 hydroxyproline-ric
390 35 77.8 164 2 G71427 hypothetical prote
391 35 77.8 166 2 T09593 cIC protein, cold-
392 35 77.8 168 2 T20708 hypothetical prote
393 35 77.8 169 2 T07623 extensin homolog H
394 35 77.8 172 2 A39458 carotene biosynthe

395	35	77.8	176	2	B85355	hypothetical prote	468	35	77.8	385	2	S78100	MAPK-activated pro
396	35	77.8	181	2	S14974	extensin class I (469	35	77.8	387	2	C96720	hypothetical prote
397	35	77.8	182	2	T30760	hypothetical prote	470	35	77.8	388	2	S25298	extensin (clone To
398	35	77.8	190	2	T35570	hypothetical prote	471	35	77.8	388	2	G85147	hypothetical prote
399	35	77.8	192	2	T30477	hypothetical prote	472	35	77.8	389	2	D87460	ribonuclease D [lm
400	35	77.8	199	2	T07622	extensin homolog -	473	35	77.8	391	2	S69192	serine O-acetyltra
401	35	77.8	209	2	C89005	protein T24A6.3 [i	474	35	77.8	392	2	B48423	homeotic protein e
402	35	77.8	214	2	A12371	hypothetical prote	475	35	77.8	395	2	I49575	CCAAT/enhancer bin
403	35	77.8	214	2	A47715	cardiac-specific h	476	35	77.8	396	2	S39793	MAPK-activated pro
404	35	77.8	216	2	T20186	hypothetical prote	477	35	77.8	398	2	JQ1245	hypothetical 43K p
405	35	77.8	218	2	F84748	probable AP2 domai	478	35	77.8	399	2	I49754	homeobox protein -
406	35	77.8	233	2	T17218	hypothetical prote	479	35	77.8	401	2	A48423	engrailed homeodom
407	35	77.8	249	2	T41847	AcMNPV orf106 - Bo	480	35	77.8	404	1	QQB313	BMRF1 protein - hu
408	35	77.8	250	2	D86214	hypothetical prote	481	35	77.8	408	2	G96707	hypothetical prote
409	35	77.8	250	2	T08071	L-ascorbate peroxi	482	35	77.8	414	2	T50010	hypothetical prote
410	35	77.8	250	2	S43157	L-ascorbate peroxi	483	35	77.8	416	2	G86232	cysteine proteinas
411	35	77.8	252	2	B48725	MDV specific prote	484	35	77.8	419	2	S56073	opaque-2 protein -
412	35	77.8	263	2	S20866	L-ascorbate peroxi	485	35	77.8	421	1	S11674	acrosin (EC 3.4.21
413	35	77.8	264	2	PQ0478	pistil extensin-II	486	35	77.8	423	2	T48121	hypothetical prote
414	35	77.8	264	2	T10572	hypothetical prote	487	35	77.8	427	2	B96804	nucellin-like prot
415	35	77.8	265	2	T04834	hypothetical prote	488	35	77.8	432	2	T06782	extensin - soybean
416	35	77.8	265	2	T46089	proline-rich prote	489	35	77.8	432	2	A43448	thrombin receptor
417	35	77.8	267	2	S08314	cell wall glycopro	490	35	77.8	432	2	B96515	hypothetical prote
418	35	77.8	267	2	T15645	hypothetical prote	491	35	77.8	433	2	T07910	hydroxyproline-ric
419	35	77.8	268	2	H84684	En/Spm-like transp	492	35	77.8	434	2	B84684	hypothetical prote
420	35	77.8	275	2	T02559	probable spliceoso	493	35	77.8	436	2	T22253	hypothetical prote
421	35	77.8	280	2	T11671	extensin-like prot	494	35	77.8	438	1	XXMSN	phosphatidylcholin
422	35	77.8	281	2	D70845	hypothetical prote	495	35	77.8	438	2	T12494	hypothetical prote
423	35	77.8	283	2	G01926	insulin promoter f	496	35	77.8	439	2	B86302	hypothetical prote
424	35	77.8	284	2	S39581	IPF1 protein - mou	497	35	77.8	440	2	I49681	glyceraldehyde-3-p
425	35	77.8	289	2	A43562	homeotic protein H	498	35	77.8	440	2	S51614	Algal-CAM - Volvox
426	35	77.8	290	2	T07080	shock protein SRC2	499	35	77.8	442	2	A38592	retinoic acid rece
427	35	77.8	296	2	T43336	csb3 protein - fis	500	35	77.8	443	1	C35991	retinoic acid rece
428	35	77.8	300	2	T08453	hypothetical prote	501	35	77.8	443	1	S29334	transcription fact
429	35	77.8	302	2	A96661	unknown protein, 8	502	35	77.8	444	2	I51256	retinoic acid rece
430	35	77.8	303	2	S28264	hydroxyproline-ric	503	35	77.8	445	1	A49447	transcription fact
431	35	77.8	310	2	T45873	hypothetical prote	504	35	77.8	445	1	S31224	transcription fact
432	35	77.8	311	2	T03827	myb protein homolo	505	35	77.8	445	2	H96560	hypothetical prote
433	35	77.8	311	2	T02783	probable homeotic	506	35	77.8	446	2	A34418	H-2 region II bind
434	35	77.8	314	2	T48514	hypothetical prote	507	35	77.8	446	2	T07907	hydroxyproline-ric
435	35	77.8	318	2	T29479	hypothetical prote	508	35	77.8	447	2	B34714	retinoic acid rece
436	35	77.8	320	2	T18319	hypothetical prote	509	35	77.8	447	2	T49439	hypothetical prote
437	35	77.8	322	2	S25299	extensin precursor	510	35	77.8	448	2	D41727	retinoid X recepto
438	35	77.8	322	2	T22403	hypothetical prote	511	35	77.8	451	2	A41651	retinoic acid rece
439	35	77.8	325	2	T31474	hypothetical prote	512	35	77.8	454	1	A33903	retinoic acid rece
440	35	77.8	325	2	A55558	albumin D-box bind	513	35	77.8	454	2	S06124	retinoic acid rece
441	35	77.8	326	2	A45452	transcription fact	514	35	77.8	458	2	A34714	retinoic acid rece
442	35	77.8	328	2	JQ0985	hydroxyproline-ric	515	35	77.8	459	2	T26878	hypothetical prote
443	35	77.8	329	2	E70973	probable lytB prot	516	35	77.8	459	2	S03116	gene 33 protein, h
444	35	77.8	330	2	T05717	probable extensin	517	35	77.8	461	2	T10741	extensin-like prot
445	35	77.8	331	2	E96529	hypothetical prote	518	35	77.8	467	2	T34874	hypothetical prote
446	35	77.8	338	2	H96765	hypothetical prote	519	35	77.8	473	2	B85187	glycoprotein homol
447	35	77.8	341	2	T37502	hypothetical prote	520	35	77.8	475	2	T03170	hypothetical prote
448	35	77.8	345	2	T01348	hypothetical prote	521	35	77.8	476	2	B41977	retinoic acid rece
449	35	77.8	350	2	S22456	hydroxyproline-ric	522	35	77.8	476	2	T27051	hypothetical prote
450	35	77.8	350	2	G75571	MutT/nudix family	523	35	77.8	477	2	T46304	hypothetical prote
451	35	77.8	352	2	T33664	hypothetical prote	524	35	77.8	478	2	JC4940	synapsin IIB - hum
452	35	77.8	357	2	T47367	hypothetical prote	525	35	77.8	479	1	S15031	paired box transcr
453	35	77.8	358	2	JC4311	CCAAT/enhancer bin	526	35	77.8	485	2	A33647	sulfated surface g
454	35	77.8	358	2	A54265	CCAAT/enhancer-bin	527	35	77.8	485	2	PC4427	Nck, Ash and phosp
455	35	77.8	358	2	A75561	conserved hypothet	528	35	77.8	487	1	S52261	NADH2 dehydrogenas
456	35	77.8	365	2	S47657	peroxidase (EC 1.1	529	35	77.8	487	2	S42442	nuclear protein EB
457	35	77.8	365	2	A34894	hepatic transcript	530	35	77.8	489	2	T11622	extensin class 1 p
458	35	77.8	368	2	C29356	hydroxyproline-ric	531	35	77.8	494	2	B96534	hypothetical prote
459	35	77.8	368	2	T51200	hypothetical prote	532	35	77.8	499	2	I51257	retinoic acid rece
460	35	77.8	370	2	JC2204	MAPK-activated pro	533	35	77.8	502	2	A55197	Wiskott-Aldrich sy
461	35	77.8	375	2	S58484	gag protein - maiz	534	35	77.8	520	2	I84718	RXR-beta1 isoform
462	35	77.8	376	2	S45763	hypothetical prote	535	35	77.8	520	2	E97813	WASP, N-WASP, MENA
463	35	77.8	377	2	JC4368	protein farnesyltr	536	35	77.8	520	2	G86414	probable protein k
464	35	77.8	377	2	A41625	protein farnesyltr	537	35	77.8	520	2	S78502	paired box transcr
465	35	77.8	379	2	S42529	Opaque-2-related p	538	35	77.8	521	2	S54266	glycoprotein gC -
466	35	77.8	379	2	S31719	proline-rich prote	539	35	77.8	529	1	W7AD22	early E2A DNA-bind
467	35	77.8	381	2	T27806	hypothetical prote	540	35	77.8	529	1	W7AD25	early E2A DNA-bind

541	35	77.8	530	2	A45690	transactivator EBN	614	35	77.8	981	2	A41401	mineralocorticoid
542	35	77.8	532	2	G84775	probable E2F5 fami	615	35	77.8	984	2	A29513	mineralocorticoid
543	35	77.8	533	2	S37781	retinoid x recepto	616	35	77.8	1001	2	T28897	hypothetical prote
544	35	77.8	534	1	A48529	ubiquinol-cytochro	617	35	77.8	1015	2	T42013	frequency clock pr
545	35	77.8	540	2	S44830	F54F2.5 protein -	618	35	77.8	1032	2	D83637	serine/threonine p
546	35	77.8	542	2	A44358	zyxin - chicken	619	35	77.8	1039	2	T22117	hypothetical prote
547	35	77.8	543	1	ERADDG	fiber protein - ca	620	35	77.8	1040	2	T29092	TSC-22 protein hom
548	35	77.8	548	2	S52735	CW17R protein - mo	621	35	77.8	1047	2	A55617	masquerade precurs
549	35	77.8	550	2	G70597	probable proteinas	622	35	77.8	1091	2	T13170	diaphanous protein
550	35	77.8	554	2	S75969	hypothetical prote	623	35	77.8	1092	2	T18305	replication factor
551	35	77.8	557	2	S62522	nuclear protein SP	624	35	77.8	1092	2	T18306	replication factor
552	35	77.8	561	2	T23722	hypothetical prote	625	35	77.8	1100	2	T30967	transcription acti
553	35	77.8	568	2	JC7317	cyclin-dependent k	626	35	77.8	1118	2	A48292	mucin, tracheobron
554	35	77.8	571	2	G84426	hypothetical prote	627	35	77.8	1119	2	T16720	hypothetical prote
555	35	77.8	572	2	G02845	zyxin - human	628	35	77.8	1173	2	T31421	C-terminal domain-
556	35	77.8	575	2	C88346	protein F42G4.3a [629	35	77.8	1176	2	T49482	hypothetical prote
557	35	77.8	580	2	T10863	extensin precursor	630	35	77.8	1188	2	S49915	extensin-like prot
558	35	77.8	581	2	E86408	F3H9.11 protein -	631	35	77.8	1209	2	T52523	hypothetical prote
559	35	77.8	585	2	I58403	H4 protein - human	632	35	77.8	1213	2	A41724	limb deformity (ld
560	35	77.8	599	2	T10798	pherophorin-S - Vo	633	35	77.8	1252	2	T14272	cortactin-binding
561	35	77.8	603	2	T22111	hypothetical prote	634	35	77.8	1276	2	E96776	hypothetical prote
562	35	77.8	605	1	QQUE3R	BVRF2 (EC-RF3) pro	635	35	77.8	1296	2	T13936	collar protein iso
563	35	77.8	620	2	S06733	hydroxyproline-ric	636	35	77.8	1297	2	S25714	son-of-sevenless-2
564	35	77.8	633	2	F84564	probable protein k	637	35	77.8	1306	2	T13592	hypothetical prote
565	35	77.8	639	2	T13151	adaptor protein CM	638	35	77.8	1366	2	B84924	hypothetical prote
566	35	77.8	639	2	G02919	transcription fact	639	35	77.8	1390	2	T31353	polyprotein - Arab
567	35	77.8	639	2	H86362	hypothetical prote	640	35	77.8	1392	2	T51947	probable transcrip
568	35	77.8	644	2	S15464	gp70 protein - mur	641	35	77.8	1400	2	T52359	hypothetical prote
569	35	77.8	653	2	JC8013	Ran-binding protei	642	35	77.8	1429	2	T13720	probable transcrip
570	35	77.8	658	2	T08153	cysteine proteinas	643	35	77.8	1584	2	T00026	hypothetical prote
571	35	77.8	664	2	C84747	probable protein k	644	35	77.8	1621	2	T15264	gene expanded prot
572	35	77.8	664	2	T01368	hypothetical prote	645	35	77.8	1736	2	A47747	brain-specific ang
573	35	77.8	666	2	T31461	probable magnesium	646	35	77.8	1878	2	E86189	hypothetical prote
574	35	77.8	678	2	AF1841	Mg chelatase chain	647	35	77.8	1905	2	T18267	multidrug resistan
575	35	77.8	694	2	T01005	hypothetical prote	648	35	77.8	1952	2	T48814	hypothetical prote
576	35	77.8	700	2	S09699	bib protein - frui	649	35	77.8	1953	2	S63244	BNi1 protein - yea
577	35	77.8	707	2	S60588	drebrin A - rat	650	35	77.8	2109	2	T17490	polyketide synthas
578	35	77.8	711	2	S43464	ecdysteroid-induce	651	35	77.8	2363	2	T38841	probable pre-mRNA
579	35	77.8	716	2	G01627	androgen receptor	652	35	77.8	2500	2	G88493	protein F57B9.2 [i
580	35	77.8	731	2	B86369	hypothetical prote	653	35	77.8	2555	2	A40043	notch protein homo
581	35	77.8	734	2	T04876	hypothetical prote	654	35	77.8	2606	2	T03159	large tegument pro
582	35	77.8	736	2	JH0681	gephyrin - rat	655	35	77.8	2706	2	T28155	variant-specific s
583	35	77.8	737	2	S28030	DNA-binding protei	656	35	77.8	2810	2	T22298	hypothetical prote
584	35	77.8	744	1	QQUEA7	UL69 protein - hum	657	35	77.8	3119	2	I49729	HD protein - mouse
585	35	77.8	744	2	E86255	hypothetical prote	658	35	77.8	3164	1	WMBEH6	UL36 protein - hum
586	35	77.8	745	2	T51370	hypothetical prote	659	34.5	76.7	407	2	T24951	hypothetical prote
587	35	77.8	754	2	T06249	protoporphyrin IX	660	34.5	76.7	409	1	VCVW2S	env polyprotein (v
588	35	77.8	755	2	T47731	hypothetical prote	661	34.5	76.7	437	2	B82778	cell cycle protein
589	35	77.8	758	2	S60586	glucocorticoid rec	662	34.5	76.7	886	2	A54442	3',5'-cyclic-nucle
590	35	77.8	758	2	T02925	protoporphyrin IX	663	34	75.6	10	2	A36454	trypsin-modulating
591	35	77.8	760	2	B96724	hypothetical prote	664	34	75.6	15	2	PT0037	light harvesting c
592	35	77.8	762	2	C96653	hypothetical prote	665	34	75.6	52	2	F71353	hypothetical prote
593	35	77.8	764	2	A56208	DNA-binding protei	666	34	75.6	57	2	S10782	salivary protein p
594	35	77.8	772	2	T13078	KIAA0992 protein -	667	34	75.6	65	2	T06457	outer envelope mem
595	35	77.8	778	2	B86218	protein T27G7.20 [668	34	75.6	66	2	T31844	hypothetical prote
596	35	77.8	782	2	A82940	hypothetical prote	669	34	75.6	74	2	B40513	hypothetical prote
597	35	77.8	786	2	T01456	extensin homolog F	670	34	75.6	74	2	T33086	hypothetical prote
598	35	77.8	789	2	T52067	hypothetical prote	671	34	75.6	76	2	C38355	basic proline-rich
599	35	77.8	791	2	C82940	hypothetical prote	672	34	75.6	77	2	T15349	hypothetical prote
600	35	77.8	805	2	T41810	ACMNPV orf66 - Bom	673	34	75.6	79	1	PJHUSB	proline-rich pepti
601	35	77.8	808	2	C72858	AcOrf-66 protein -	674	34	75.6	79	2	E87607	hypothetical prote
602	35	77.8	811	2	S04085	ovarian tumor prot	675	34	75.6	81	2	T30482	hypothetical prote
603	35	77.8	817	2	T03852	protein phosphatas	676	34	75.6	90	2	T18068	hypothetical prote
604	35	77.8	820	2	T00645	hypothetical prote	677	34	75.6	91	2	T11554	vpv protein - simi
605	35	77.8	845	2	T17291	hypothetical prote	678	34	75.6	91	2	T30512	hypothetical prote
606	35	77.8	846	2	T21700	hypothetical prote	679	34	75.6	92	2	T34146	hypothetical prote
607	35	77.8	847	2	F96531	hypothetical prote	680	34	75.6	103	2	T29597	hypothetical prote
608	35	77.8	884	2	D96730	unknown protein F5	681	34	75.6	104	2	G84607	hypothetical prote
609	35	77.8	899	2	A35895	androgen receptor	682	34	75.6	105	2	T22564	hypothetical prote
610	35	77.8	902	2	B40494	androgen receptor	683	34	75.6	106	2	B72490	hypothetical prote
611	35	77.8	958	2	T13593	hypothetical prote	684	34	75.6	108	2	T26880	hypothetical prote
612	35	77.8	962	2	T00262	hypothetical prote	685	34	75.6	108	2	A48831	vitelline membrane
613	35	77.8	975	2	T48107	hypothetical prote	686	34	75.6	108	2	F72653	hypothetical prote

687	34	75.6	111	2	S08438	vpx protein - huma	760	34	75.6	188	2	C96593	unknown protein, 9
688	34	75.6	112	1	ASLJGH	vpu protein - huma	761	34	75.6	189	2	T10580	zinc-finger protei
689	34	75.6	112	1	ASLJSX	vpu protein - huma	762	34	75.6	189	2	E90769	EspF-like protein
690	34	75.6	112	2	T11562	vpx protein - simi	763	34	75.6	191	2	T28682	hypothetical prote
691	34	75.6	112	2	S07990	vpx protein - simi	764	34	75.6	194	2	A38203	proline-rich prote
692	34	75.6	112	2	S57448	DNA binding protei	765	34	75.6	195	2	T07735	nodulin-20a - soyb
693	34	75.6	115	2	T36886	hypothetical prote	766	34	75.6	195	2	T21300	hypothetical prote
694	34	75.6	117	2	T46322	hypothetical prote	767	34	75.6	195	2	I52578	ALL-1 protein - hu
695	34	75.6	121	2	S17718	anther-specific pr	768	34	75.6	196	2	B44525	Ca2+-transporting
696	34	75.6	121	2	S12245	anther-specific pr	769	34	75.6	197	2	S35252	proline-rich prote
697	34	75.6	129	2	T19263	hypothetical prote	770	34	75.6	198	2	E86261	F13K23.6 protein -
698	34	75.6	133	2	T26755	hypothetical prote	771	34	75.6	202	2	T11744	dehydrin - kidney
699	34	75.6	133	2	H87681	hypothetical prote	772	34	75.6	204	2	T07679	protein import rec
700	34	75.6	134	2	D84672	hypothetical prote	773	34	75.6	205	2	C44525	Ca2+-transporting
701	34	75.6	134	2	T36365	proline-rich prote	774	34	75.6	205	2	F87623	hypothetical prote
702	34	75.6	135	2	E96750	hypothetical prote	775	34	75.6	207	2	I53154	scleraxis - mouse
703	34	75.6	137	2	T22308	hypothetical prote	776	34	75.6	209	2	T07779	dehydrin homolog C
704	34	75.6	139	2	T37932	very hypothetical	777	34	75.6	211	2	B89716	protein F45B8.3 [i
705	34	75.6	139	2	A71099	hypothetical prote	778	34	75.6	212	2	S74288	hypothetical prote
706	34	75.6	139	2	C87459	hypothetical prote	779	34	75.6	213	2	T48490	embryo-specific pr
707	34	75.6	139	2	S61885	hypothetical prote	780	34	75.6	214	2	S65052	pistil-specific pr
708	34	75.6	139	4	S09612	extensin precursor	781	34	75.6	215	2	T22572	hypothetical prote
709	34	75.6	141	2	T06646	blue copper-bindin	782	34	75.6	217	2	T09965	extensin CYC17 pre
710	34	75.6	141	2	A34043	hypothetical proli	783	34	75.6	217	2	T15873	hypothetical prote
711	34	75.6	142	2	T51039	related to extensi	784	34	75.6	217	4	I65366	ALL-1/AF-4 mutant
712	34	75.6	143	2	S42579	QID3 protein - fun	785	34	75.6	218	2	T23318	hypothetical prote
713	34	75.6	143	2	E72699	hypothetical prote	786	34	75.6	220	2	S52005	nodulin-30 (Npv30)
714	34	75.6	145	2	E88462	protein C05H8.2 [i	787	34	75.6	220	2	S42879	nodulin-30 - kidne
715	34	75.6	145	2	T48552	glutaredoxin-like	788	34	75.6	221	2	H82857	TonB protein XF000
716	34	75.6	148	2	S46956	outer membrane pro	789	34	75.6	222	2	T43500	hypothetical prote
717	34	75.6	148	2	S46542	outer envelope pro	790	34	75.6	223	2	A42817	proline-rich prote
718	34	75.6	150	2	C72697	hypothetical prote	791	34	75.6	224	2	G87377	hypothetical prote
719	34	75.6	151	2	S10084	hypothetical 16K p	792	34	75.6	225	2	T09964	extensin CYC15 pre
720	34	75.6	153	2	T31654	hypothetical prote	793	34	75.6	225	2	D88492	protein T07E3.6 [i
721	34	75.6	154	2	T41831	AcMNPV orf91 - Bom	794	34	75.6	226	2	S41032	hypothetical prote
722	34	75.6	155	2	T16855	hypothetical prote	795	34	75.6	226	2	T35236	hypothetical prote
723	34	75.6	156	2	T29730	hypothetical prote	796	34	75.6	230	2	T36672	membrane-spanning
724	34	75.6	157	2	T15064	hypothetical prote	797	34	75.6	231	2	T50853	response regulator
725	34	75.6	158	2	A86452	protein F6N18.7 [i	798	34	75.6	231	2	S37108	cuticlin 2 - Caeno
726	34	75.6	159	2	D72486	hypothetical prote	799	34	75.6	235	2	E75342	conserved hypothet
727	34	75.6	160	2	S58759	ezzrin - rat (fragm	800	34	75.6	236	2	T45835	hypothetical prote
728	34	75.6	160	2	AC2958	conserved hypothet	801	34	75.6	236	2	T02577	probable AP2 domai
729	34	75.6	161	2	T45055	hypothetical prote	802	34	75.6	239	2	T03078	conserved hypothet
730	34	75.6	163	2	D84674	hypothetical prote	803	34	75.6	240	2	T25814	hypothetical prote
731	34	75.6	163	2	T33130	hypothetical prote	804	34	75.6	240	2	D70894	probable pra prote
732	34	75.6	164	2	C98325	probable transcrip	805	34	75.6	241	2	T22216	hypothetical prote
733	34	75.6	165	2	C71717	hypothetical prote	806	34	75.6	243	2	S28444	hypothetical prote
734	34	75.6	165	2	T24470	hypothetical prote	807	34	75.6	244	2	B87417	tonB protein - Pae
735	34	75.6	166	1	PIHUSC	salivary proline-r	808	34	75.6	244	2	S44822	hypothetical prote
736	34	75.6	166	2	B25372	salivary proline-r	809	34	75.6	246	2	T46446	F44E2.3 protein -
737	34	75.6	167	2	A44525	Ca2+-transporting	810	34	75.6	247	2	T06721	hypothetical prote
738	34	75.6	169	2	T18321	hypothetical prote	811	34	75.6	247	2	T50874	hypothetical cytoc
739	34	75.6	171	2	A27307	proline-rich phosp	812	34	75.6	247	2	T33469	hypothetical prote
740	34	75.6	172	2	S68232	antimicrobial prot	813	34	75.6	248	2	F83558	hypothetical prote
741	34	75.6	172	2	B29149	proline-rich prote	814	34	75.6	249	2	A41497	36K antigen pra -
742	34	75.6	172	2	D41132	collagen-related p	815	34	75.6	250	1	A31757	homeotic protein H
743	34	75.6	172	2	T27505	hypothetical prote	816	34	75.6	250	2	E81734	conserved hypothet
744	34	75.6	173	2	T51469	glycine/proline-ri	817	34	75.6	250	2	T08908	hypothetical prote
745	34	75.6	174	2	A41358	gag polyprotein -	818	34	75.6	250	2	T16342	hypothetical prote
746	34	75.6	174	2	B96543	hypothetical prote	819	34	75.6	250	2	A85632	hypothetical prote
747	34	75.6	175	2	S71560	early light-induce	820	34	75.6	251	1	B60492	homeotic protein H
748	34	75.6	180	2	E84768	hypothetical prote	821	34	75.6	253	2	T32465	hypothetical prote
749	34	75.6	180	2	D72629	hypothetical prote	822	34	75.6	254	2	S52236	MADS box protein a
750	34	75.6	181	2	JC5233	spermatophorin Sp2	823	34	75.6	254	2	B84901	hypothetical prote
751	34	75.6	181	2	T48558	hypothetical prote	824	34	75.6	254	2	AE2516	hypothetical prote
752	34	75.6	182	2	A44157	spermatophorin Sp2	825	34	75.6	255	2	C36222	phosphoprotein pho
753	34	75.6	182	2	B35650	hypothetical prote	826	34	75.6	255	2	T34163	hypothetical prote
754	34	75.6	183	2	JC4356	ankyrin precursor	827	34	75.6	256	2	S39794	MAPK-activated pro
755	34	75.6	184	2	S78091	endocuticular prot	828	34	75.6	262	2	S23241	hypothetical prote
756	34	75.6	185	2	B26669	nodulin-20 precurs	829	34	75.6	263	2	T12698	extensin - common
757	34	75.6	185	2	B84531	hypothetical prote	830	34	75.6	264	2	B89005	protein T24A6.18 l
758	34	75.6	186	2	C41132	collagen-related p	831	34	75.6	268	2	F70823	hypothetical prote
759	34	75.6	187	2	S38036	hypothetical prote	832	34	75.6	269	2	C84707	hypothetical prote

833 34 75.6 269 2 T30468 hypothetical prote
834 34 75.6 269 2 A71401 hypothetical prote
835 34 75.6 270 2 A84549 probable C2H2-type
836 34 75.6 270 2 S63049 hypothetical prote
837 34 75.6 271 2 T26640 hypothetical prote
838 34 75.6 271 2 T49956 hypothetical prote
839 34 75.6 273 2 F87414 phage SPO1 DNA pol
840 34 75.6 274 2 T04619 hypothetical prote
841 34 75.6 275 2 S75818 hypothetical prote
842 34 75.6 277 2 H87552 hypothetical prote
843 34 75.6 277 2 S78063 homeobox protein p
844 34 75.6 279 2 T02495 hypothetical prote
845 34 75.6 280 2 A53027 transcription fact
846 34 75.6 281 2 T29150 hypothetical prote
847 34 75.6 281 2 S36748 transcriptional prote
848 34 75.6 282 2 T26635 hypothetical prote
849 34 75.6 282 2 A53336 transcription fact
850 34 75.6 283 2 S13383 hydroxyproline-ric
851 34 75.6 287 2 T34397 hypothetical prote
852 34 75.6 288 2 T45715 hypothetical prote
853 34 75.6 289 2 T52354 hypothetical prote
854 34 75.6 289 2 AE2006 hypothetical prote
855 34 75.6 290 2 AC2030 hypothetical prote
856 34 75.6 291 2 T27534 hypothetical prote
857 34 75.6 294 2 A55477 survival motor neu
858 34 75.6 297 2 B83049 conserved hypothet
859 34 75.6 301 2 T18788 hypothetical prote
860 34 75.6 301 2 T50001 hypothetical prote
861 34 75.6 302 2 S31865 chlorophyll a/b-bi
862 34 75.6 303 1 TVHUJD transforming prote
863 34 75.6 304 2 A32993 transcription fact
864 34 75.6 304 2 T48281 hypothetical prote
865 34 75.6 304 2 H86332 T20H2.26 protein -
866 34 75.6 308 2 A86246 hypothetical prote
867 34 75.6 309 1 S09257 homeotic protein H
868 34 75.6 309 2 E36596 nuclear factor I (
869 34 75.6 309 2 T47685 probable RNA bindi
870 34 75.6 311 2 S20085 transcription fact
871 34 75.6 312 2 D87475 rare lipoprotein A
872 34 75.6 312 2 T29232 hypothetical prote
873 34 75.6 313 1 FOVDA gag polypeptide -
874 34 75.6 314 2 D71410 hypothetical prote
875 34 75.6 315 2 JQ1250 coat protein - chr
876 34 75.6 315 2 T34932 hypothetical prote
877 34 75.6 315 2 A64689 hypothetical prote
878 34 75.6 315 2 A56561 35K proline-rich p
879 34 75.6 323 2 S20099 transforming prote
880 34 75.6 323 2 H70765 hypothetical prote
881 34 75.6 324 2 S13497 cAMP-binding prote
882 34 75.6 324 2 T24102 hypothetical prote
883 34 75.6 328 2 C42756 hypothetical prote
884 34 75.6 329 2 T10064 cytokinin-induced
885 34 75.6 330 2 S37595 mucin JUL10 - huma
886 34 75.6 331 2 B47236 zinc-finger protei
887 34 75.6 333 2 S11484 CABP1 protein - sl
888 34 75.6 334 2 T04572 hypothetical prote
889 34 75.6 337 2 C90968 EspF-like protei
890 34 75.6 338 1 TVMSFB transforming prote
891 34 75.6 338 2 I53043 transforming prote
892 34 75.6 338 2 I56893 transcription fact
893 34 75.6 340 2 S53763 homeotic protein H
894 34 75.6 344 1 AHMS Ig alpha chain C r
895 34 75.6 344 2 A86294 hypothetical prote
896 34 75.6 345 2 S12788 transcription fact
897 34 75.6 345 2 A48462 dense granule prot
898 34 75.6 345 2 C86272 F7A19.1 protein -
899 34 75.6 346 2 S19129 proline-rich prote
900 34 75.6 346 2 S76923 hypothetical prote
901 34 75.6 347 2 A43815 transforming prote
902 34 75.6 348 2 A49594 enhancer factor pr
903 34 75.6 348 2 AB3260 hypothetical membr
904 34 75.6 348 2 S52720 homeobox protein g
905 34 75.6 349 2 G84852 hypothetical prote

34 75.6 350 2 T14191 extensin homolog T
34 75.6 352 2 C96643 hypothetical prote
34 75.6 354 1 GNVVSR genome polypeptide
34 75.6 355 2 S41285 coat protein - swe
34 75.6 356 1 WJHU2H homeotic protein H
34 75.6 357 2 PC4293 nuclear factor 1 f
34 75.6 358 2 T10244 G-Box binding prot
34 75.6 358 2 E87309 hypothetical prote
34 75.6 358 2 T01296 leucine-rich repea
34 75.6 359 2 T13478 hypothetical prote
34 75.6 360 2 D87332 hypothetical prote
34 75.6 364 2 I48188 gene Nkx6.1 protei
34 75.6 365 2 T04718 hypothetical prote
34 75.6 365 2 T24955 hypothetical prote
34 75.6 365 2 B42832 factor VIII intron
34 75.6 366 2 T30457 hypothetical prote
34 75.6 366 2 I53035 trithorax homolog
34 75.6 367 1 S02193 cellular tumor ant
34 75.6 368 2 T03828 myb protein - rice
34 75.6 370 2 B88455 protein T15B12.2 [
34 75.6 370 2 JC6130 paired box transcr
34 75.6 372 2 S74859 hypothetical prote
34 75.6 372 2 T10472 G-box binding prot
34 75.6 373 2 E84595 hypothetical prote
34 75.6 373 2 A70856 probable lppz prot
34 75.6 376 2 A26066 segmentation prote
34 75.6 378 2 T21651 hypothetical prote
34 75.6 379 2 S49999 40K protein - frog
34 75.6 381 2 G02668 neurogenic basic-h
34 75.6 381 2 S59679 hypothetical prote
34 75.6 381 2 B86175 protein F19P19.21
34 75.6 382 2 E85082 hypothetical prote
34 75.6 382 2 T14186 hypothetical prote
34 75.6 383 2 T00674 hypothetical prote
34 75.6 384 2 H88924 protein F33E11.2 [
34 75.6 384 2 S51796 vasodilator-stimul
34 75.6 384 2 C85816 unknown protein en
34 75.6 385 2 T20410 hypothetical prote
34 75.6 388 2 T26922 hypothetical prote
34 75.6 389 2 C70921 probable molybdopt
34 75.6 389 2 T03411 protein RFL - rice
34 75.6 389 2 S27200 proline-rich prote
34 75.6 390 2 B82819 outer membrane pro
34 75.6 390 2 G01936 Abl binding protei
34 75.6 392 2 F84650 probable protein p
34 75.6 393 2 T33103 lin-1 protein - Ca
34 75.6 393 2 JC5614 RNB6 protein - rat
34 75.6 394 2 C84905 probable extensin
34 75.6 396 1 A58938 surface protein rh
34 75.6 396 2 T07021 extensin-like prot
34 75.6 396 2 A36339 FLO protein - gard
34 75.6 396 2 S41224 hnRNP protein - Af
34 75.6 396 2 S62964 hypothetical prote
34 75.6 396 2 S75266 hypothetical prote
34 75.6 398 2 AI3231 conjugal transfer
34 75.6 399 2 T01345 hypothetical prote
34 75.6 404 2 H88650 protein C09G12.9 [
34 75.6 405 2 T42663 hypothetical prote
34 75.6 406 2 T36632 probable oxidoredu
34 75.6 407 1 A40150 inhibin beta-B cha
34 75.6 410 2 C70889 hypothetical prote
34 75.6 411 2 B41398 inhibin beta-B cha
34 75.6 413 2 T03240 FLO/LFY protein ho
34 75.6 413 2 H87604 hypothetical prote
34 75.6 414 2 C86301 arginine/serine-ri
34 75.6 415 1 A34170 acrosin (EC 3.4.21
34 75.6 416 2 T03243 FLO/LFY protein ho
34 75.6 417 2 T15174 hypothetical prote
34 75.6 418 2 F70940 hypothetical prote
34 75.6 418 2 T16713 hypothetical prote
34 75.6 419 2 T19871 hypothetical prote
34 75.6 420 2 T46910 hypothetical prote
34 75.6 420 2 JC4716 zinc finger DNA-bi

979 34 75.6 427 2 T48159 hypothetical prote
980 34 75.6 427 2 I49603 transcription regu
981 34 75.6 428 1 TVHUEK transforming prote
982 34 75.6 428 2 E71415 probable coll wall
983 34 75.6 429 2 T06296 extensin-like prot
984 34 75.6 429 2 A47305 translation initia
985 34 75.6 430 2 JC2301 hypothetical 47.8K
986 34 75.6 431 2 S47538 acrosin (EC 3.4.21
987 34 75.6 431 2 T12450 hypothetical prote
988 34 75.6 433 1 FOLJH2 gag polypeptid -
989 34 75.6 434 2 C96515 hypothetical prote
990 34 75.6 435 2 A42672 choline kinase (EC
991 34 75.6 436 2 B55452 cartilage-derived
992 34 75.6 439 2 AF1048 N-acetylmuramoyl-L
993 34 75.6 440 2 S48953 hypothetical prote
994 34 75.6 440 2 H86312 F2H15.2 protein -
995 34 75.6 440 2 JC7807 Wiskott-Aldrich sy
996 34 75.6 440 2 S71795 transcription fact
997 34 75.6 441 1 TVCHMC transforming prote
998 34 75.6 441 2 T22531 hypothetical prote
999 34 75.6 442 2 T34018 hypothetical prote
1000 34 75.6 443 2 G86388 unknown protein [i

ALIGNMENTS

RESULT 1
A53692
synapsin I - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C:Accession: A53692
R;Chin, L.S.; Li, L.; Greengard, P.
J. Biol. Chem. 269, 18507-18513, 1994
A:Title: Neuron-specific expression of the synapsin II gene is directed by a specific co
A:Reference number: A53692; MUID:94308086; PMID:8034599
A:Accession: A53692
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-125 <RES>
A:Cross-references: UNIPROT:O88935; UNIPARC:UPI000016D062; GB:L32025; NID:G1041084; PIDN:
C:Genetics:
A:Gene: SYN1

Query Match 100.0%; Score 45; DB 2; Length 125;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 27 RPQPPPP 33

RESULT 2
A25704
synapsin I - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 21-May-1988 #sequence_revision 21-May-1988 #text_change 09-Jul-2004
C:Accession: A25704
R;McCaffery, C.A.; DeGennaro, L.J.
EMBO J. 5, 3167-3173, 1986
A:Title: Determination and analysis of the primary structure of the nerve terminal speci
A:Reference number: A25704; MUID:87133474; PMID:3028773
A:Accession: A25704
A:Molecule type: mRNA
A:Residues: 1-691 <MCC>
A:Cross-references: UNIPROT:P09951; UNIPARC:UPI0000170B12; GB:X04655; NID:G57181; PIDN:
A:Experimental source: brain
C:Keywords: actin binding; alternative splicing; phosphoprotein

.Query Match 100.0%; Score 45; DB 2; Length 691;
Best Local Similarity 100.0%; Pred. No. 60;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RPQPPPP 7
Db 14 RPQPPPP 20
RESULT 3
A30411
synapsin Ia - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 23-Oct-1990 #sequence_revision 23-Oct-1990 #text_change 09-Jul-2004
C:Accession: A30411; B30411
R;Suedhof, T.C.; Czernik, A.J.; Kao, H.T.; Takei, K.; Johnston, P.A.; Horiuchi, A.; Kana
Science 245, 1474-1480, 1989
A:Title: Synapsins: mosaics of shared and individual domains in a family of synaptic vesic
A:Reference number: A30411; MUID:89388265; PMID:2506642
A:Accession: A30411
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-704 <SUE>
A:Cross-references: UNIPROT:P09951; UNIPARC:UPI000004368E; GB:M27812; NID:G206920; PIDN:
A:Accession: B30411
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-659, 'KA', 662, 'PAQAQP' <SU2>
A:Cross-references: UNIPARC:UPI000002B48E; GB:M27924; NID:G206932; PIDN:AAA42148.1; PID:
C:Keywords: actin binding; alternative splicing; phosphoprotein

Query Match 100.0%; Score 45; DB 2; Length 704;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 27 RPQPPPP 33

RESULT 4
A35363
synapsin I splice form a - human
C:Species: Homo sapiens (man)
C:Date: 17-Aug-1990 #sequence_revision 17-Aug-1990 #text_change 09-Jul-2004
C:Accession: A35363; B35363; A35805
R;Suedhof, T.C.
J. Biol. Chem. 265, 7849-7852, 1990
A:Title: The structure of the human synapsin I gene and protein.
A:Reference number: A35363; MUID:90243651; PMID:2110562
A:Accession: A35363
A:Molecule type: DNA
A:Residues: 1-705 <SUE>
A:Cross-references: UNIPROT:P17600; UNIPARC:UPI0000179AF3; GB:M58371; GB:J05431
A:Accession: B35363
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-659, 'KASPAQAQP' <SU2>
A:Cross-references: UNIPARC:UPI0000179AF4; GB:M58378; GB:J05431
R;Sauerwald, A.; Hoesche, C.; Oschwald, R.; Kilimann, M.W.
J. Biol. Chem. 265, 14932-14937, 1990
A:Title: The 5'-flanking region of the synapsin I gene. A G+C-rich, TATA- and CAAT-less,
A:Reference number: A35805; MUID:90368667; PMID:2118519
A:Accession: A35805
A:Molecule type: DNA
A:Residues: 1-125 <SAU>
A:Cross-references: UNIPARC:UPI000016B07E; GB:M55301; NID:G338655; PIDN:AAA60608.1; PID:
C:Genetics:
A:Gene: GDB:SYN1
A:Cross-references: GDB:119606; OMIM:313440
A:Map position: Xp11.23-Xp11.23
C:Keywords: actin binding; alternative splicing; phosphoprotein

Query Match 100.0%; Score 45; DB 2; Length 705;
Best Local Similarity 100.0%; Pred. No. 61;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
Db 27 RPQPPPP 33

RESULT 5
E30411
synapsin Ia - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 23-Oct-1990 #sequence_revision 23-Oct-1990 #text_change 01-Dec-2000
C;Accession: E30411; F30411; A35758
R;Suedhof, T.C.; Czernik, A.J.; Kao, H.T.; Takei, K.; Johnston, P.A.; Horiuchi, A.; Kana
Science 245, 1474-1480, 1989
A;Title: Synapsins: mosaics of shared and individual domains in a family of synaptic ves
A;Reference number: A30411; MUID:89388265; PMID:2506642
A;Accession: E30411
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-706 <SUE>
A;Cross-references: UNIPARC:UPI0000179AF5
A;Accession: F30411
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-661,'KA',664,'PAQAQP' <SU2>
A;Cross-references: UNIPARC:UPI0000179AF5
R;Hall, F.L.; Mitchell, J.P.; Vulliet, P.R.
J. Biol. Chem. 265, 6944-6948, 1990
A;Title: Phosphorylation of synapsin I at a novel site by proline-directed protein kinas
A;Reference number: A35758; MUID:90216728; PMID:2108963
A;Accession: A35758
A;Molecule type: protein
A;Residues: 532-556 <HAL>
A;Cross-references: UNIPARC:UPI0000179AF7
C;Keywords: actin binding; alternative splicing; phosphoprotein
F;551/Binding site: phosphate (Ser) (covalent) (by proline-directed kinase) #status expe

Query Match 100.0%; Score 45; DB 2; Length 706;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
Db 27 RPQPPPP 33

RESULT 6
T03455
ALR protein - human
C;Species: Homo sapiens (man)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C;Accession: T03455
R;Prasad, R.; Zhadanov, A.B.; Sedkov, Y.; Bullrich, F.; Druck, T.; Rallapalli, R.; Yano,
Oncogene 15, 549-560, 1997
A;Title: Structure and expression pattern of human ALR, a novel gene with strong homolog
A;Reference number: Z14954; MUID:97388474; PMID:9247308
A;Accession: T03455
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-4957 <PRA>
A;Cross-references: UNIPROT:O14686; UNIPARC:UPI00001100F1; EMBL:AF010404; NID:g23358286;
C;Genetics:
A;Gene: ALR
A;Map position: 12
C;Superfamily: acute lymphoblastic leukemia protein, ALR type
C;Keywords: alternative splicing

Query Match 100.0%; Score 45; DB 2; Length 4957;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7

Db 1806 RPQPPPP 1812

RESULT 7
T03454
ALR protein - human
C;Species: Homo sapiens (man)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C;Accession: T03454
R;Prasad, R.; Zhadanov, A.B.; Sedkov, Y.; Bullrich, F.; Druck, T.; Rallapalli, R.; Yano,
Oncogene 15, 549-560, 1997
A;Title: Structure and expression pattern of human ALR, a novel gene with strong homolog
A;Reference number: Z14954; MUID:97388474; PMID:9247308
A;Accession: T03454
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-5262 <PRA>
A;Cross-references: UNIPROT:O14686; UNIPARC:UPI0000110107; EMBL:AF010403; NID:g23358284;
C;Genetics:
A;Gene: ALR
A;Map position: 12
C;Superfamily: acute lymphoblastic leukemia protein, ALR type
C;Keywords: alternative splicing

Query Match 100.0%; Score 45; DB 2; Length 5262;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
Db 2111 RPQPPPP 2117

RESULT 8
T05847
hypothetical protein F17L22.180 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T05847
R;Bevan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voet, M.; Robben, J.; Volckaert, G.; Bar
submitted to the Protein Sequence Database, February 1999
A;Reference number: Z15454
A;Accession: T05847
A;Molecule type: DNA
A;Residues: 1-139 <BEV>
A;Cross-references: UNIPROT:Q9SVS5; UNIPARC:UPI00000AASA4; EMBL:AL035527
A;Experimental source: cultivar Columbia; BAC clone F17L22
C;Genetics:
A;Map position: 4
A;Introns: 26/1; 44/1; 78/3; 91/3
A;Note: F17L22.180

Query Match 93.3%; Score 42; DB 2; Length 139;
Best Local Similarity 85.7%; Pred. No. 31;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
Db 113 KPQPPPP 119

RESULT 9
AE2842
conserved hypothetical protein Atu2166 [imported] - Agrobacterium tumefaciens (strain C58
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AE2842
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E

ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AE2842
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-320 <KUR>
A;Cross-references: UNIPROT:Q8UDG0; UNIPARC:UPI00001646C6; GB:AE008688; PIDN:AAL43155.1;
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu2166
A;Map position: circular chromosome

Query Match 93.3%; Score 42; DB 2; Length 320;
Best Local Similarity 85.7%; Pred. No. 69;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
:|||||
Db 74 KPQPPPP 80

RESULT 10
F97619
hypothetical protein AGR_C_3928 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: F97619
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: F97619
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <KUR>
A;Cross-references: UNIPROT:Q8UDG0; UNIPARC:UPI00000D1DDF; GB:AE007869; PIDN:AAK87911.1;
C;Genetics:
A;Gene: AGR_C_3928
A;Map position: circular chromosome

Query Match 93.3%; Score 42; DB 2; Length 359;
Best Local Similarity 85.7%; Pred. No. 77;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
:|||||
Db 113 KPQPPPP 119

RESULT 11
A71416
hypothetical protein - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
A;Variety: columbia
C;Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 09-Jul-2004
C;Accession: A71416
R;Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenegger, T.; Pohl, T.M.; Terry, N.; Giel
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A;Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech
erthoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ana
C.; Chalwatzis, N.
A;Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal
A;Reference number: A71400; MUID:98121113; PMID:9461215
A;Accession: A71416
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-645 <BEV>
A;Cross-references: UNIPROT:O23374; UNIPARC:UPI00000A2BC7; GB:Z97338; NID:g2244870; PID:
C;Genetics:

A;Map position: 4COP9-4G3845

Query Match 93.3%; Score 42; DB 2; Length 645;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
:|||||
Db 329 KPQPPPP 335

RESULT 12
I51691
dishevelled homolog - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C;Accession: I51691
R;Sokol, S.Y.; Klingensmith, J.; Perrimon, N.; Itoh, K.
Development 121, 1637-1647, 1995
A;Title: Dorsalizing and neuralizing properties of Xdsh, a maternally expressed Xenopus
A;Reference number: I51691; MUID:95324391; PMID:7600981
A;Accession: I51691
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-736 <SOK>
A;Cross-references: UNIPROT:P51142; UNIPARC:UPI00001299A8; EMBL:U31552; NID:g945109; PID:
C;Genetics:
A;Gene: Xdsh
F;260-336/Domain: GLGF domain homology <GLG4>

Query Match 93.3%; Score 42; DB 2; Length 736;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
:|||||
Db 98 RPEPPPP 104

RESULT 13
T30188
Niemann-Pick C disease protein - mouse
N;Alternate names: NCP1 protein
C;Species: Mus musculus (house mouse)
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C;Accession: T30188
R;Loftus, S.K.; Morris, J.A.; Carstea, E.D.; Gu, J.Z.; Cummings, C.; Brown, A.; Ellison,
Science 277, 232-235, 1997
A;Title: Murine model of niemann-pick C disease: mutation in a cholesterol homeostasis g
A;Reference number: Z20765; MUID:97362324; PMID:9211850
A;Accession: T30188
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1278 <LOF>
A;Cross-references: UNIPROT:O35604; UNIPARC:UPI00000299BC; EMBL:AF003348; NID:g2251241; I
C;Genetics:
A;Gene: Npc1
A;Map position: 18

Query Match 93.3%; Score 42; DB 2; Length 1278;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
:|||||
Db 251 KPQPPPP 257

RESULT 14
A46431
tight junction-associated protein ZO-1 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004

C;Accession: A46431
R;Itoh, M.; Nagafuchi, A.; Yonemura, S.; Kitani-Yasuda, T.; Tsukita, S.; Tsukita, S.
J. Cell Biol. 121, 491-502, 1993
A;Title: The 220-kD protein colocalizing with cadherins in non-epithelial cells is identical.
A;Reference number: A46431; MUID:93252986; PMID:8486731
A;Accession: A46431
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-1745 <ITO>
A;Cross-references: UNIPROT:P39447; UNIPARC:UPI0000029DB4; GB:D14340; NID:G303709; PIDN:
A;Experimental source: F9 cells
A;Note: sequence extracted from NCBI backbone (NCBIN:131200, NCBIP:131201)
F;27-106/Domain: GLGF domain homology <GLG1>
F;428-498/Domain: GLGF domain homology <GLG3>
F;645-794/Domain: guanylate kinase homology <GKI>

Query Match 93.3%; Score 42; DB 2; Length 1745;
Best Local Similarity 85.7%; Pred. No. 3.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
:|||||
Db 1241 KPQPPPP 1247

RESULT 15
A83295
hypothetical protein PA2819 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: A83295
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrner, P.; Hickey, M.J.; Bha
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: A83295
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-159 <STO>
A;Cross-references: UNIPROT:Q9I024; UNIPARC:UPI00000C5825; GB:AE004708; GB:AE004091; NID
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA2819

Query Match 91.1%; Score 41; DB 2; Length 159;
Best Local Similarity 85.7%; Pred. No. 47;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
:|||||
Db 50 RPRPPPP 56

RESULT 16
JQ1682
infected cell protein ICP34.5 - human herpesvirus 1 (strain 17)
C;Species: human herpesvirus 1
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004
C;Accession: JQ1682
R;Dolan, A.; McKie, E.; MacLean, A.R.; McGeoch, D.J.
J. Gen. Virol. 73, 971-973, 1992
A;Title: Status of the ICP34.5 gene in herpes simplex virus type 1 strain 17.
A;Reference number: JQ1682; MUID:92341080; PMID:1321882
A;Accession: JQ1682
A;Molecule type: DNA
A;Residues: 1-248 <DOL>
A;Cross-references: UNIPROT:P36313; UNIPARC:UPI000012D17E; GB:S40593
C;Genetics:
A;Gene: ICP34.5
C;Superfamily: herpesvirus infected cell protein ICP34.5

C;Keywords: tandem repeat
F;161-175/Region: 3-residue repeats (A-T-P)

Query Match 91.1%; Score 41; DB 1; Length 248;
Best Local Similarity 85.7%; Pred. No. 73;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
:|||||
Db 102 RPRPPPP 108

RESULT 17
A39724
homeotic protein Hox A4 - human
N;Alternate names: homeotic protein Hox 1.4; homeotic protein Hox 1D
C;Species: Homo sapiens (man)
C;Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 09-Jul-2004
C;Accession: A39724; A60492; S15544; A32468
R;Buettnner, R.; Yim, S.O.; Hong, Y.S.; Boncinelli, E.; Tainsky, M.A.
Mol. Cell. Biol. 11, 3573-3583, 1991
A;Title: Alteration of homeobox gene expression by N-ras transformation of PA-1 human te
A;Reference number: A39724; MUID:91260707; PMID:1675427
A;Accession: A39724
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-320 <BUE>
A;Cross-references: UNIPROT:Q00056; UNIPARC:UPI000016AA7D; GB:M39724
R;Peverali, F.A.; D'Esposito, M.; Acampora, D.; Bunone, G.; Negri, M.; Faiella, A.; Stor
Differantiation 45, 61-69, 1990
A;Title: Expression of HOX homeogenes in human neuroblastoma cell culture lines.
A;Reference number: A60492; MUID:91153613; PMID:1981366
A;Accession: A60492
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-107,'A',109-320 <PEV>
A;Cross-references: UNIPARC:UPI000012CF25
R;Boncinelli, E.; Acampora, D.; Pannese, M.; d'Esposito, M.; Somma, R.; Gaudino, G.; Sto
Genome 31, 745-756, 1989
A;Title: Organization of human class I homeobox genes.
A;Reference number: S15036; MUID:90215256; PMID:2576652
A;Accession: S15544
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 215-280 <BON>
A;Cross-references: UNIPARC:UPI00001745B6
R;Ferguson-Smith, A.C.; Fienberg, A.; Ruddie, F.H.
Genomics 5, 250-258, 1989
A;Title: Isolation, chromosomal localization, and nucleotide sequence of the human HOX 1
A;Reference number: A32468; MUID:90007544; PMID:2571574
A;Accession: A32468
A;Molecule type: DNA
A;Residues: 215-275 <FER>
A;Cross-references: UNIPARC:UPI00001745B7; EMBL:M28199
C;Comment: This homeotic protein is expressed in embryonic nervous system.
C;Genetics:
A;Gene: GDB:HOXA4
A;Cross-references: GDB:120650; OMIM:142953
A;Map position: 7p15.3-7p15.3
C;Function:
A;Description: control of embryonic development by tissue- and stage-specific regulation
C;Superfamily: homeotic protein Hox D4; homeobox homology
C;Keywords: DNA binding; embryo; homeobox; nucleus; transcription regulation
F;216-272/Domain: homeobox homology <HOX>

Query Match 91.1%; Score 41; DB 1; Length 320;
Best Local Similarity 85.7%; Pred. No. 93;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
:|||||
Db 137 QPQPPPP 143

RESULT 18
A41786
mRNA-binding protein p54 - African clawed frog
C/Species: Xenopus laevis (African clawed frog)
C/Date: 04-Mar-1993 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C/Accession: A41786; A36348
R/Murray, M.T.; Schiller, D.L.; Franke, W.W.
Proc. Natl. Acad. Sci. U.S.A. 89, 11-15, 1992
A/Title: Sequence analysis of cytoplasmic mRNA-binding proteins of Xenopus oocytes identified
A/Reference number: A41786; MUID:92107999; PMID:1729676
A/Accession: A41786
A/Molecule type: mRNA
A/Residues: 1-324 <MUR>
A/Cross-references: UNIPROT:P45441; UNIPARC:UPI000013A23D; GB:M80257; NID:G214641; PIDN:
A/Note: sequence extracted from NCBI backbone (NCBIN:74686, NCBIP:74687)
R/Murray, M.T.; Krohne, G.; Franke, W.W.
J. Cell Biol. 112, 1-11, 1991
A/Title: Different forms of soluble cytoplasmic mRNA binding proteins and particles in X
A/Reference number: A36348; MUID:91093331; PMID:1670777
A/Accession: A36348
A/Status: nucleic acid sequence not shown; not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 221-233 <MU2>
A/Cross-references: UNIPARC:UPI00001745A4
A/Note: authors say this sequence was found in the similar p56 molecule but there are two
C/Superfamily: Xenopus Y box-binding protein 2; cold shock domain homology
C/Keywords: DNA binding; nucleus; oocyte; RNA binding; transcription regulation
F;44-108/Domain: cold shock domain homology <CSD>

Query Match 91.1%; Score 41; DB 1; Length 324;
Best Local Similarity 85.7%; Pred. No. 94;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
||:||||
Db 238 RPRPPPP 244

RESULT 19
T00388
hypothetical protein KIAA0616 - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
C/Accession: T00388
R/Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.
DNA Res. 5, 169-176, 1998
A/Title: Prediction of the coding sequences of unidentified human genes. X. The complete
A/Reference number: Z14142; MUID:98403880; PMID:9734811
A/Accession: T00388
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: mRNA
A/Residues: 1-634 <ISH>
A/Cross-references: UNIPROT:O75114; UNIPARC:UPI0000073D49; EMBL:AB014516; NID:G3327045;
A/Experimental source: brain
C/Genetics:
A/Gene: KIAA0616

Query Match 91.1%; Score 41; DB 2; Length 634;
Best Local Similarity 85.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
||:||||
Db 366 QPQPPPP 372

RESULT 20
T00366
hypothetical protein KIAA0669 - human
C/Species: Homo sapiens (man)
C/Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
C/Accession: T00366

R/Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.
DNA Res. 5, 169-176, 1998
A/Title: Prediction of the coding sequences of unidentified human genes. X. The complete
A/Reference number: Z14142; MUID:98403880; PMID:9734811
A/Accession: T00366
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: mRNA
A/Residues: 1-780 <ISH>
A/Cross-references: UNIPROT:O75157; UNIPARC:UPI00000722E0; EMBL:AB014569; NID:G3327151;
A/Experimental source: brain; clone HK02346
C/Genetics:
A/Note: KIAA0669

Query Match 91.1%; Score 41; DB 2; Length 780;
Best Local Similarity 85.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
||:||||
Db 273 QPQPPPP 279

RESULT 21
T47922
anthranilate phosphoribosyltransferase-like protein - Arabidopsis thaliana
N/Alternate names: protein T20K12.200
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C/Accession: T47922
R/De Haan, M.; Maarse, A.C.; Grivell, L.A.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quetf
submitted to the Protein Sequence Database, January 2000
A/Reference number: Z24480
A/Accession: T47922
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-972 <DEH>
A/Cross-references: UNIPROT:Q9M2D4; UNIPARC:UPI000000C1A5; EMBL:AL137898
A/Experimental source: cultivar Columbia; BAC clone T20K12
C/Genetics:
A/Map position: 3
A/Note: T20K12.200
C/Superfamily: Arabidopsis thaliana hypothetical protein F20D22.8

Query Match 91.1%; Score 41; DB 2; Length 972;
Best Local Similarity 85.7%; Pred. No. 2.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
||:||||
Db 178 QPQPPPP 184

RESULT 22
A35913
regulatory factor X - human
C/Species: Homo sapiens (man)
C/Date: 08-Mar-1991 #sequence_revision 08-Mar-1991 #text_change 09-Jul-2004
C/Accession: A35913
R/Reich, W.; Herrero-Sanchez, C.; Kobr, M.; Silacci, P.; Berte, C.; Barras, E.; Fey, S.;
Genes Dev. 4, 1528-1540, 1990
A/Title: MHC class II regulatory factor RFX has a novel DNA-binding domain and a function
A/Reference number: A35913; MUID:91071581; PMID:2253877
A/Accession: A35913
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-979 <REI>
A/Cross-references: UNIPROT:P22670; UNIPARC:UPI000002CE9D; GB:X58964; NID:G311362; PIDN:
C/Keywords: DNA binding; transcription regulation

Query Match 91.1%; Score 41; DB 2; Length 979;
Best Local Similarity 85.7%; Pred. No. 2.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
:|||||
Db 26 QPQPPPP 32

RESULT 23
T30351
mucin-like protein - Lymantria dispar nuclear polyhedrosis virus
C:Species: Lymantria dispar nuclear polyhedrosis virus, LdMNPV
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T30351
R:Kuzio, J.; Pearson, M.N.; Harwood, S.H.; Funk, C.J.; Evans, J.T.; Slavicek, J.M.; Rohd
Virology 253, 17-34, 1999
A:Title: Sequence and analysis of the genome of a baculovirus pathogenic for Lymantria d
A:Reference number: Z20836; MUID:99124785; PMID:9887315
A:Accession: T30351
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1029 <KUZ>
A:Cross-references: UNIPROT:Q9YMX0; UNIPARC:UPI00000EC0CA; EMBL:AF081810; PIDN:AAC70189.

Query Match 91.1%; Score 41; DB 2; Length 1029;
Best Local Similarity 85.7%; Pred. No. 2.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
:|||||
Db 933 QPQPPPP 939

RESULT 24
A57278
fibrillin-2 precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 09-Jul-2004
C:Accession: A57278
R:Zhang, H.; Hu, W.; Ramirez, F.
J. Cell Biol. 129, 1165-1176, 1995
A:Title: Developmental expression of fibrillin genes suggests heterogeneity of extracell
A:Reference number: A57278; MUID:95263670; PMID:7744963
A:Accession: A57278
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-2907 <ZHA>
A:Cross-references: UNIPROT:Q61555; UNIPARC:UPI00000289B9; GB:L39790; NID:g762830; PIDN:
C:Superfamily: fibrillin; EGF homology
F:1239-1274/Domain: EGF homology <EGF1>
F:2488-2523/Domain: EGF homology <EGF>

Query Match 91.1%; Score 41; DB 2; Length 2907;
Best Local Similarity 85.7%; Pred. No. 7.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
:|||||
Db 29 QPQPPPP 35

RESULT 25
A54105
fibrillin-2 precursor - human
C:Species: Homo sapiens (man)
C:Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 09-Jul-2004
C:Accession: A54105; S17063; S31101
R:Zhang, H.; Apfelroth, S.D.; Hu, W.; Davis, E.C.; Sanguinetti, C.; Bonadio, J.; Mecham,
J. Cell Biol. 124, 855-863, 1994
A:Title: Structure and expression of fibrillin-2, a novel microfibrillar component prefe
A:Reference number: A54105; MUID:94165150; PMID:8120105
A:Accession: A54105
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A:Molecule type: mRNA
A:Residues: 1-2918 <ZHA>
A:Cross-references: UNIPROT:P35556; UNIPARC:UPI000017651B; GB:U03272

R:Lee, B.; Godfrey, M.; Vitale, E.; Hori, H.; Mattei, M.G.; Sarfarazi, M.; Tsipouras, P.
Nature 352, 330-334, 1991
A:Title: Linkage of Marfan syndrome and a phenotypically related disorder to two differer
A:Reference number: S17062; MUID:91304567; PMID:1852206
A:Accession: S17063
A:Molecule type: mRNA
A:Residues: 752-1489,1791,'GS',1794-1796,'QLI',1922-1923,'LD',1926,'P',1928 <LEE>
A:Cross-references: UNIPARC:UPI000017651C; EMBL:X62009
R:Milewicz, D.M.
submitted to the EMBL Data Library, December 1992
A:Reference number: S31101
A:Accession: S31101
A:Molecule type: mRNA
A:Residues: 752-1407,'R',1409-1489,1791,'GS',1794-1796,'QLI',1922-1923,'LD',1926,'P',1928
A:Cross-references: UNIPARC:UPI000000065B; EMBL:X62009
C:Genetics:
A:Gene: GDB:FBN2
A:Cross-references: GDB:128122; OMIM:121050
A:Map position: 5q23-5q31
C:Superfamily: fibrillin; EGF homology
C:Keywords: extracellular protein
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-2918/Product: fibrillin-2 #status predicted <MAT>
F:1245-1280/Domain: EGF homology <EGF1>
F:1970-2013/Domain: EGF homology <EGF>

Query Match 91.1%; Score 41; DB 2; Length 2918;
Best Local Similarity 85.7%; Pred. No. 7.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
:|||||
Db 29 QPQPPPP 35

RESULT 26
A46068
Huntington disease-associated protein - human
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: A46068; I54337
R:MacDonald, M.E.; Ambrose, C.M.; Duyao, M.P.; Myers, R.H.; Lin, C.; Srinidhi, L.; Barne
J.F.; Bates, G.P.; Baxendale, S.; Hummerich, H.; Kirby, S.; North, M.; Youngman, S.; Mott
O'Donovan, M.C.; Riba-Ramirez, L.; Shah, M.; Stanton, V.P.; Strobel, S.A.; Draths, K.M.
Cell 72, 971-983, 1993
A:Authors: Wales, J.L.; Dervan, P.; Housman, D.E.; Altherr, M.; Shiang, R.; Thompson, L.;
d, K.; Collins, F.S.; Snell, R.; Holloway, T.; Gillespie, K.; Datson, N.; Shaw, D.; Harpe
A:Title: A novel gene containing a trinucleotide repeat that is expanded and unstable on
A:Reference number: A46068; MUID:93208892; PMID:8458085
A:Accession: A46068
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-3144 <MAC>
A:Cross-references: UNIPROT:P42858; UNIPARC:UPI0000146F05; GB:L12392
R:Lin, B.; Rommens, J.M.; Graham, R.K.; Kalchman, M.; MacDonald, H.; Nasir, J.; Delaney,
Hum. Mol. Genet. 2, 1541-1545, 1993
A:Title: Differential 3' polyadenylation of the Huntington disease gene results in two m
A:Reference number: I54337; MUID:94093536; PMID:7903579
A:Accession: I54337
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 2563-3144 <RES>
A:Cross-references: UNIPARC:UPI000016AABF; GB:L20431; NID:g398028; PIDN:AAA52702.1; PID:
C:Genetics:
A:Gene: GDB:HD
A:Cross-references: GDB:119307; OMIM:143100
A:Map position: 4p16.3-4p16.3

Query Match 91.1%; Score 41; DB 2; Length 3144;
Best Local Similarity 85.7%; Pred. No. 8.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7

Db :|||||
66 QPQPPPP 72

RESULT 27
A96650
protein F24O1.6 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: A96650
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: A96650
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-70 <STO>
A;Cross-references: UNIPROT:Q9MAV4; UNIPARC:UPI00000A2B99; GB:AE005173; NID:g7940276; PID:
C;Genetics:
A;Gene: F24O1.6
A;Map position: 1
C;Superfamily: hydroxyproline-rich glycoprotein

Query Match 88.9%; Score 40; DB 2; Length 70;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PQPPPP 7
|||
Db 35 PQPPPP 40

RESULT 28
B48232
cysteine-rich extensin-like protein 2 precursor - common tobacco
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 09-Jul-2004
C;Accession: B48232; PQ0474; S24616
R;Wu, H.; Zou, J.; May, B.; Gu, Q.; Cheung, A.Y.
Proc. Natl. Acad. Sci. U.S.A. 90, 6829-6833, 1993
A;Title: A tobacco gene family for flower cell wall proteins with a proline-rich domain
A;Reference number: A48232; MUID:933005740; PMID:8341705
A;Accession: B48232
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-196 <WUA>
A;Cross-references: UNIPROT:Q08195; UNIPARC:UPI00000A3E44; GB:L13440; NID:g310924; PIDN:
Ride S Goldman, M.H.; Pezzotti, M.; Seurinck, J.; Mariani, C.
Plant Cell 4, 1041-1051, 1992
A;Title: Developmental expression of tobacco pistil-specific genes encoding novel extensi
A;Reference number: PQ0474; MUID:933005740; PMID:1392607
A;Accession: PQ0474
A;Molecule type: mRNA
A;Residues: 'MAG', 1-105 <GOL>
A;Cross-references: UNIPARC:UPI0000177E6E; EMBL:Z14014
A;Experimental source: stigma, style; strain Petit Havana SR1
C;Genetics:
A;Gene: CELP-2
C;Superfamily: glutelin
C;Keywords: cell wall; extracellular matrix; fertilization
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-196/Product: cysteine-rich extensin-like protein 2 #status experimental <MAT>

Query Match 88.9%; Score 40; DB 2; Length 196;
Best Local Similarity 85.7%; Pred. No. 78;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
|||
Db 100 RPSPPPP 106

RESULT 29
A48232
cysteine-rich extensin-like protein 1 precursor - common tobacco
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 09-Jul-2004
C;Accession: A48232; PQ0475; S24617
R;Wu, H.; Zou, J.; May, B.; Gu, Q.; Cheung, A.Y.
Proc. Natl. Acad. Sci. U.S.A. 90, 6829-6833, 1993
A;Title: A tobacco gene family for flower cell wall proteins with a proline-rich domain
A;Reference number: A48232; MUID:93342083; PMID:8341705
A;Accession: A48232
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-209 <WUA>
A;Cross-references: UNIPROT:Q08194; UNIPARC:UPI00000ABB3F; GB:L13439; NID:g310922; PIDN:
Ride S Goldman, M.H.; Pezzotti, M.; Seurinck, J.; Mariani, C.
Plant Cell 4, 1041-1051, 1992
A;Title: Developmental expression of tobacco pistil-specific genes encoding novel extensi
A;Reference number: PQ0474; MUID:93005740; PMID:1392607
A;Accession: PQ0475
A;Molecule type: mRNA
A;Residues: 39-209 <GOL>
A;Cross-references: UNIPARC:UPI000016DE8A; EMBL:Z14020; NID:g19918; PID:g19919
A;Experimental source: stigma, style; strain Petit Havana SR1
C;Superfamily: glutelin
C;Keywords: cell wall; extracellular matrix; fertilization; glycoprotein
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-209/Product: cysteine-rich extensin-like protein 1 #status experimental <MAT>
F;146/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 88.9%; Score 40; DB 2; Length 209;
Best Local Similarity 85.7%; Pred. No. 83;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
|||
Db 115 RPSPPPP 121

RESULT 30
S34163
homeotic protein Hoxd-3 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C;Accession: S34163
R;Bedford, M.; Petr, L.
submitted to the EMBL Data Library, June 1993
A;Reference number: S34163
A;Accession: S34163
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-215 <BED>
A;Cross-references: UNIPROT:Q61685; UNIPARC:UPI00000028C7F; EMBL:X73572; NID:g313248; PID:
C;Genetics:
A;Introns: 165/3

Query Match 88.9%; Score 40; DB 2; Length 215;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PQPPPP 7
|||
Db 101 PQPPPP 106

RESULT 31

T07176
extensin homolog - potato (fragment)
C;Species: Solanum tuberosum (potato)
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
C;Accession: T07176
R;Macleod, M.R.
submitted to the EMBL Data Library, December 1997
A;Reference number: Z15977
A;Accession: T07176
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-221 <MAC>
A;Cross-references: UNIPROT:Q49946; UNIPARC:UPI00000ACA27; EMBL:AJ003220; NID:e1251331;
A;Experimental source: cv. Record; swelling stolon
C;Superfamily: glutelin

Query Match 88.9%; Score 40; DB 2; Length 221;
Best Local Similarity 85.7%; Pred. No. 87;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
|||
Db 50 RPSPPPP 56

RESULT 32
S54156
extensin-like protein - cowpea (fragment)
C;Species: Vigna unguiculata (cowpea)
C;Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
C;Accession: S54156
R;Arsenijevic-Maksimovic, I.; Broughton, W.J.; Krause, A.
Submitted to the EMBL Data Library, April 1995
A;Description: A class of root-hair specific extensins involved in rhizobium/legume inte
A;Reference number: S54155
A;Accession: S54156
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-242 <ARS>
A;Cross-references: UNIPROT:Q43687; UNIPARC:UPI000009E1E7; EMBL:X86029; NID:g791147; PID
C;Superfamily: hydroxyproline-rich glycoprotein

Query Match 88.9%; Score 40; DB 2; Length 242;
Best Local Similarity 85.7%; Pred. No. 95;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
|||
Db 128 RPSPPPP 134

RESULT 33
C86480
33.2K hypothetical protein - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: C86480
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: C86480
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-302 <STO>
A;Cross-references: UNIPROT:Q9C8B9; UNIPARC:UPI00000A8A47; GB:AE005172; NID:g11055830; H

C;Genetics:
A;Map position: 1

Query Match 88.9%; Score 40; DB 2; Length 302;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PQPPPP 7
|||||
Db 82 PQPPPP 87

RESULT 34
JQ1386
hypothetical 33K protein - pea enation mosaic virus
C;Species: pea enation mosaic virus, PEMV
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C;Accession: JQ1386
R;Demler, S.A.; de Zoeten, G.A.
J. Gen. Virol. 72, 1819-1834, 1991
A;Title: The nucleotide sequence and luteovirus-like nature of RNA 1 of an aphid non-tra
A;Reference number: JQ1382; MUID:91341468; PMID:1875194
A;Accession: JQ1386
A;Molecule type: genomic RNA
A;Residues: 1-303 <DEM>
A;Cross-references: UNIPROT:Q84711; UNIPARC:UPI00000F0B4D; GB:L04573; NID:g294105; PIDN:
A;Experimental source: strain WSG
A;Note: this reading frame extends between two stop codons and does not begin with a sta
A;Note: 65-Asp was also found
C;Genetics:
A;Map position: segment RNAL

Query Match 88.9%; Score 40; DB 2; Length 303;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PQPPPP 7
|||||
Db 14 PQPPPP 19

RESULT 35
B29350
env polypotein, retrovirus-related - mouse (fragment)
N;Contains: coat protein gp70
C;Species: Mus musculus (house mouse)
C;Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 15-Nov-1996
C;Accession: B29350
R;Levy, D.E.; Lerner, R.A.; Wilson, M.C.
J. Virol. 56, 691-700, 1985
A;Title: Normal expression of polymorphic endogenous retroviral RNA containing segments
A;Reference number: A29350; MUID:86062897; PMID:2415714
A;Accession: B29350
A;Molecule type: mRNA
A;Residues: 1-312 <LEV>
A;Cross-references: UNIPARC:UPI000017863E
A;Experimental source: clone E1
A;Note: the authors translated the codon CTA for residue 147 as Lys
C;Superfamily: type C retrovirus env polypotein
C;Keywords: coat protein; glycoprotein; polypotein; spike protein; transmembrane protei
F;22-312/Product: coat protein gp70 (fragment) #status predicted <GPS>
F;43,58,297/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 88.9%; Score 40; DB 2; Length 312;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PQPPPP 7
|||||
Db 258 PQPPPP 263

RESULT 36

E70519
Probable oxidoreductase - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: E70519
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A;Authors: Squares, R.; Suleston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: E70519
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-334 <COL>
A;Cross-references: UNIPROT:O07721; UNIPARC:UPI000004B72C; GB:Z97193; GB:AL123456; NID:9
A;Experimental source: strain H37RV
C;Genetics:
A;Gene: fadB5
C;Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology
F;26-323/Domain: long-chain alcohol dehydrogenase homology <LADH>

Query Match 88.9%; Score 40; DB 2; Length 334;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
|||
Db 20 RPDPPPP 26

RESULT 37
T33836
hypothetical protein F54A5.3a - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T33836
R;Jones, K.; Graves, T.; Ozersky, P.; Wilson, R.
submitted to the EMBL Data Library, November 1998
A;Description: The sequence of C. elegans cosmid F54A5.
A;Reference number: Z21420
A;Accession: T33836
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-334 <JON>
A;Cross-references: UNIPROT:Q9TYT3; UNIPARC:UPI000017BA0D; EMBL:AF106584; PIDN:AAC78222.
A;Experimental source: strain Bristol N2; clone F54A5
C;Genetics:
A;Gene: CESP:F54A5.3a
A;Map position: 1
A;Introns: 22/3; 46/2; 70/3; 234/2

Query Match 88.9%; Score 40; DB 2; Length 334;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
|||
Db 266 RPSPPPP 272

RESULT 38
JC5833
transcription factor E2F-5 - human
N;Alternate names: p130/p107-binding protein E2F-5
C;Species: Homo sapiens (man)
C;Date: 20-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 05-Oct-2004
C;Accession: JC5833; I38878
R;Vaishnav, Y.N.; Vaishnav, M.Y.; Pant, V.
Biochem. Biophys. Res. Commun. 242, 586-592, 1998
A;Title: The molecular and functional characterization of E2F-5 transcription factor.
A;Reference number: JC5833; MUID:98125536; PMID:9464260

A;Accession: JC5833
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-346 <VAI>
A;Cross-references: UNIPROT:Q15329; UNIPARC:UPI000002E057; EMBL:Z78409; NID:g1592737
R;Sardet, C.; Vidal, M.; Cobrinik, D.; Geng, Y.; Onufryk, C.; Chen, A.; Weinberg, R.A. Proc. Natl. Acad. Sci. U.S.A. 92, 2403-2407, 1995
A;Title: E2F-4 and E2F-5, two members of the E2F family, are expressed in the early phase
A;Reference number: I38877; MUID:95199358; PMID:7892279
A;Accession: I38878
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-294,296-346 <SAR>
A;Cross-references: UNIPARC:UPI000016A135; EMBL:U15642; NID:g758415; PIDN:AAC50120.1; PII
C;Genetics:
A;Gene: GDB:E2F5
A;Cross-references: GDB:578742; OMIM:600967
A;Map position: 16q22.1-16q22.1
C;Superfamily: transcription factor E2F5
C;Keywords: cell cycle control; DNA binding
F;49-116/Domain: DNA-binding #status predicted <DNA>

Query Match 88.9%; Score 40; DB 2; Length 346;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PQPPPP 7
|||
Db 33 PQPPPP 38

RESULT 39
T49791
hypothetical protein B9J10.280 [imported] - Neurospora crassa
C;Species: Neurospora crassa
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C;Accession: T49791
R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura, submitted to the Protein Sequence Database, May 2000
A;Reference number: Z25022
A;Accession: T49791
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-349 <SCH>
A;Cross-references: UNIPROT:Q9P555; UNIPARC:UPI000006B94E; EMBL:AL356324; GSPDB:GN00116;
A;Experimental source: BAC clone B9J10; strain OR74A
C;Genetics:
A;Gene: NCSP:B9J10.280
A;Map position: 6

Query Match 88.9%; Score 40; DB 2; Length 349;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
|||
Db 181 RPHPPPP 187

RESULT 40
VCVWB6
env polyprotein precursor - Friend spleen focus-forming virus (strain BB6)
N;Alternate names: coat polyprotein
N;Contains: knob protein gp70; spike protein p15E
C;Species: Friend spleen focus-forming virus
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 09-Jul-2004
C;Accession: A41995
R;Majumdar, M.K.; Cho, C.L.; Fox, M.T.; Eckner, K.L.; Kozak, S.; Kabat, D.; Geib, R.W. J. Virol. 66, 3652-3660, 1992
A;Title: Mutations in the env gene of Friend spleen focus-forming virus overcome Fv-2(r) -
A;Reference number: A41995; MUID:92260637; PMID:1583724
A;Accession: A41995
A;Molecule type: DNA

A;Residues: 1-356 <MAJ>
A;Cross-references: UNIPROT:P31793; UNIPARC:UPI000012A007; GB:M90673
C;Genetics:
A;Gene: env
C;Superfamily: type C retrovirus env polyprotein
C;Keywords: coat protein; glycoprotein; polyprotein
F;1-33/Domain: signal sequence #status predicted <SIG>
F;34-356/Product: env polyprotein #status predicted <ENV>
F;34-330/Product: knob protein gp70 #status predicted <KPG>
F;331-356/Product: spike protein p15E #status predicted <SPP>
F;335-351/Region: hydrophobic #status predicted
F;43,58,296,328/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 88.9%; Score 40; DB 1; Length 356;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PQPPPP 7
|||
Db 257 PQPPPP 262

RESULT 41
S31428
envelope protein - spleen focus-forming virus
C;Species: spleen focus-forming virus, SFFV
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C;Accession: S31428
R;Cho, C.L.; Majumdar, M.K.; Geib, R.W.
submitted to the EMBL Data Library, May 1992
A;Reference number: S31428
A;Accession: S31428
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-356 <CHO>
A;Cross-references: UNIPROT:Q87044; UNIPARC:UPI000010210C; EMBL:X66721; NID:g61959; PIDN
C;Superfamily: type C retrovirus env polyprotein

Query Match 88.9%; Score 40; DB 2; Length 356;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PQPPPP 7
|||
Db 257 PQPPPP 262

RESULT 42
VCMVSR
env polyprotein - Rauscher spleen focus-forming virus
N;Alternate names: coat polyprotein
N;Contains: coat protein gp70; coat protein p15E
C;Species: Rauscher spleen focus-forming virus
C;Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 09-Jul-2004
C;Accession: A03988
R;Bestwick, R.K.; Boswell, B.A.; Kabat, D.
J. Virol. 51, 695-705, 1984
A;Title: Molecular cloning of biologically active Rauscher spleen focus-forming virus an
A;Reference number: A93001; MUID:84292446; PMID:6088793
A;Accession: A03988
A;Molecule type: DNA
A;Residues: 1-408 <BES>
A;Cross-references: UNIPROT:P03389; UNIPARC:UPI000012A048; GB:K02375; NID:g331998; PIDN
C;Genetics:
A;Gene: env
C;Superfamily: type C retrovirus env polyprotein
C;Keywords: coat protein; glycoprotein; polyprotein
F;1-33/Domain: leader peptide #status predicted <LDP>
F;34-333/Product: coat protein gp70 #status predicted <CPA>
F;334-408/Product: coat protein p15E #status predicted <CPE>
F;43,58,297,329,378/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 88.9%; Score 40; DB 1; Length 408;

Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 PQPPPP 7
|||
Db 258 PQPPPP 263
RESULT 43
S50899
betaB inhibin precursor - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
C;Accession: S50899
R;Thompson, D.A.; Cronin, C.N.; Martin, F.
Eur. J. Biochem. 226, 751-764, 1994
A;Title: Genomic cloning and sequence analyses of the bovine alpha-, beta(A)- and beta(B
Y DNase I footprinting.
A;Reference number: S50897; MUID:95112839; PMID:7813465
A;Accession: S50899
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-408 <THO>
A;Cross-references: UNIPROT:P42917; UNIPARC:UPI000012D425; EMBL:U16240
C;Genetics:
A;Introns: 151/1
C;Superfamily: inhibin

Query Match 88.9%; Score 40; DB 2; Length 408;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PQPPPP 7
|||
Db 39 PQPPPP 44

RESULT 44
A21170
gp55 protein - Friend spleen focus-forming virus
C;Species: Friend spleen focus-forming virus
C;Date: 14-Dec-1990 #sequence_revision 14-Dec-1990 #text_change 08-Apr-1994
C;Accession: A21170
R;Amanuma, H.; Katori, A.; Obata, M.; Sagata, N.; Ikawa, Y.
Proc. Natl. Acad. Sci. U.S.A. 80, 3913-3917, 1983
A;Title: Complete nucleotide sequence of the gene for the specific glycoprotein (gp55) o
A;Reference number: A21170; MUID:83247388; PMID:6306650
A;Accession: A21170
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-409 <AMA>
A;Cross-references: UNIPARC:UPI0000178640
C;Superfamily: type C retrovirus env polyprotein

Query Match 88.9%; Score 40; DB 2; Length 409;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PQPPPP 7
|||
Db 257 PQPPPP 262

RESULT 45
S34164
homeotic protein Hoxd-3 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Oct-2004
C;Accession: S34164
R;Bedford, M.; Petr, L.
submitted to the EMBL Data Library, June 1993
A;Reference number: S34163
A;Accession: S34164

A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-411 <BED>
A;Cross-references: UNIPROT:P09027; UNIPARC:UPI000016CE02; EMBL:X73573; NID:g313252; PID:
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;180-236/Domain: homeobox homology <HOX>

Query Match 88.9%; Score 40; DB 2; Length 411;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PQPPPP 7
| | | | |
Db 101 PQPPPP 106

RESULT 46
T04520
hypothetical protein F16A16.100 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
C;Accession: T04520
R;Bevan, M.; Brandt, P.; Dose, S.; Jarke, D.; Scharfe, M.; Schon, O.; Hoheisel, J.; Mewe
submitted to the Protein Sequence Database, February 1999
A;Reference number: Z15376
A;Accession: T04520
A;Molecule type: DNA
A;Residues: 1-413 <BEV>
A;Cross-references: UNIPROT:Q9SVU6; UNIPARC:UPI00000A6306; EMBL:AL035353
A;Experimental source: cultivar Columbia; BAC clone F16A16
C;Genetics:
A;Map position: 4
A;Introns: 22/2; 262/3; 287/3; 309/3; 331/3
A;Note: F16A16.100

Query Match 88.9%; Score 40; DB 2; Length 413;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PQPPPP 7
| | | | |
Db 69 PQPPPP 74

RESULT 47
S27198
homeotic protein Hox D3 - human
N;Alternate names: homeotic protein Hox 4A
C;Species: Homo sapiens (man)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C;Accession: S27198
R;Taniguchi, Y.; Fujii, A.; Moriuchi, T.
Biochim. Biophys. Acta 1132, 332-334, 1992
A;Title: Cloning and sequencing of the human homeobox gene HOX4A.
A;Reference number: S27198; MUID:93041940; PMID:1358204
A;Accession: S27198
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-416 <TAN>
A;Cross-references: UNIPROT:P31249; UNIPARC:UPI000016AA89; EMBL:D11117; NID:g219879; PID:
C;Genetics:
A;Introns: 165/1
C;Superfamily: homeotic protein Hox B3; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;179-235/Domain: homeobox homology <HOX>

Query Match 88.9%; Score 40; DB 2; Length 416;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PQPPPP 7
| | | | |
Db 101 PQPPPP 106

RESULT 48
T34279
hypothetical protein F46H5.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T34279
R;Nhan, M.
submitted to the EMBL Data Library, November 1995
A;Description: The sequence of C. elegans cosmid F46H5.
A;Reference number: Z21498
A;Accession: T34279
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-416 <NHA>
A;Cross-references: UNIPROT:P90878; UNIPARC:UPI000007BC53; EMBL:U41543; PIDN:AAB37027.1;
A;Experimental source: strain Bristol N2; clone F46H5
C;Genetics:
A;Gene: CESP:F46H5.6
A;Map position: X
A;Introns: 231/3; 260/3; 325/3; 352/2; 385/2

Query Match 88.9%; Score 40; DB 2; Length 416;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PQPPPP 7
| | | | |
Db 39 PQPPPP 44

RESULT 49
S47539
homeotic protein Hox D3 - mouse
N;Alternate names: homeobox-containing protein Hox-4.1
C;Species: Mus musculus (house mouse)
C;Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 09-Jul-2004
C;Accession: S47539; JC6144
R;Brown, W.M.; Zhou, L.; Taylor, G.R.
Biochim. Biophys. Acta 1219, 219-222, 1994
A;Title: The nucleotide sequence of the murine Hox-D3 (Hox-4.1) gene reveals extensive i
A;Reference number: S47539; MUID:94368862; PMID:7916214
A;Accession: S47539
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-417 <BRO>
A;Cross-references: UNIPROT:P09027; UNIPARC:UPI0000026946; EMBL:U03496
R;Tan, D.P.; Shao, X.; Pu, L.; Guo, V.; Nirenberg, M.
Proc. Natl. Acad. Sci. U.S.A. 93, 8247-8252, 1996
A;Title: Sequence and expression of the murine Hoxd-3 homeobox gene.
A;Reference number: JC6144; MUID:96323206; PMID:8710855
A;Contents: embryo
A;Accession: JC6144
A;Molecule type: mRNA
A;Residues: 1-417 <TAN>
A;Cross-references: UNIPARC:UPI0000026946; GB:U56400; NID:g1398928; PIDN:AAC52779.1; PID:
C;Comment: This protein is a homeodomain protein with SH3 domain, and it controls the rat
C;Genetics:
A;Gene: Hoxd-3
A;Introns: 166/1
C;Superfamily: homeotic protein Hox B3; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;180-236/Domain: homeobox homology <HOX>

Query Match 88.9%; Score 40; DB 2; Length 417;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PQPPPP 7
| | | | |
Db 101 PQPPPP 106

RESULT 50
A83184
probable protein methyltransferase PA3706 [imported] - Pseudomonas aeruginosa (strain PA
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: A83184
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: A83184
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-422 <STO>
A;Cross-references: UNIPROT:Q9HXT5; UNIPARC:UPI00000C5ACD; GB:AE004789; GB:AE004091; NID
C;Experimental source: strain PA01
C;Genetics:
A;Gene: PA3706

Query Match 88.9%; Score 40; DB 2; Length 422;
Best Local Similarity 85.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
|||
Db 292 RPSPPPP 298

Search completed: April 6, 2006, 09:40:09
Job time : 17.7105 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2006, 09:23:24 ; Search time 85.1053 Seconds
(without alignments)
58.030 Million cell updates/sec

Title: US-10-632-388-305
Perfect score: 45
Sequence: 1 RQPPPPP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : UniProt 05.80: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	100.0	66	2	Q6IIP9 DROME
2	45	100.0	73	2	Q5VPI6 ORYSA
3	45	100.0	119	2	Q5VPA7 ORYSA
4	45	100.0	125	2	Q6LDK2 9MURI
5	45	100.0	132	2	Q6YZB3 ORYSA
6	45	100.0	165	2	Q7F8E9 ORYSA
7	45	100.0	208	2	Q654U7 ORYSA
8	45	100.0	270	2	Q9ZA63 NEIME
9	45	100.0	291	2	Q6Z7W0 ORYSA
10	45	100.0	300	2	Q8LIW2 ORYSA
11	45	100.0	317	2	Q6MPB5 BDEBA
12	45	100.0	328	2	Q4T9Z6 TETNG
13	45	100.0	398	2	Q9LHG8 ARATH
14	45	100.0	416	2	Q9C2D6 NEUCR
15	45	100.0	420	2	Q9VZC2 DROME
16	45	100.0	449	2	Q9PZ18 GVXN
17	45	100.0	551	2	Q4POX1 USTMA
18	45	100.0	587	2	Q525T5 MAGGR
19	45	100.0	624	2	Q4TA45 TETNG
20	45	100.0	669	2	Q5H9B0 HUMAN
21	45	100.0	687	2	Q4G0U1 HUMAN
22	45	100.0	704	1	SYN1_RAT
23	45	100.0	704	2	Q95XQ7 CABEL
24	45	100.0	705	1	SYN1_HUMAN
25	45	100.0	706	1	SYN1_BOVIN
26	45	100.0	706	1	SYN1_MOUSE
27	45	100.0	721	2	Q51L36 MAGGR
28	45	100.0	1412	2	Q4QF52 LEIMA
29	45	100.0	1480	2	Q96Q04 HUMAN
30	45	100.0	1798	1	MAST2_HUMAN
31	45	100.0	2301	2	Q9ATK5_CHLRE
					Q6IIP9 drosophila
					Q5VP16 oryza sativ
					Q5VPA7 oryza sativ
					Q6LDK2 rattus sp.
					Q6YZB3 oryza sativ
					Q7F8E9 oryza sativ
					Q654U7 oryza sativ
					Q9ZA63 neisseria m
					Q6Z7W0 oryza sativ
					Q8LIW2 oryza sativ
					Q6MPB5 bdellovibri
					Q4T9Z6 tetraodon n
					Q9LHG8 arabidopsis
					Q9C2D6 neurospora
					Q9VZC2 drosophila
					Q9PZ18 xestia c-ni
					Q4POX1 ustilago ma
					Q525T5 magnaporthe
					Q4TA45 tetraodon n
					Q5H9B0 homo sapien
					Q4G0U1 homo sapien
					P09951 rattus norv
					Q95XQ7 caenorhabdi
					P17599 bos taurus
					O88935 mus musculu
					Q51L36 magnaporthe
					Q4QF52 leishmania
					Q96Q04 homo sapien
					Q6P0Q8 homo sapien
					Q9ATK5 chlamydomon

32	45	100.0	5262	1	MLL2 HUMAN	014686 homo sapien
33	42	93.3	107	2	Q67YU2 ARATH	Q67YU2 arabidopsis
34	42	93.3	139	2	Q9SVS5 ARATH	Q9SVS5 arabidopsis
35	42	93.3	169	2	Q5N912 ORYSA	Q5N912 oryza sativ
36	42	93.3	246	2	Q6ZOS6 ORYSA	Q6ZOS6 oryza sativ
37	42	93.3	279	2	Q6ZQX5 HUMAN	Q6ZQX5 homo sapien
38	42	93.3	287	2	Q96KT3 HUMAN	Q96KT3 homo sapien
39	42	93.3	309	2	Q7XPK1 ORYSA	Q7XPK1 oryza sativ
40	42	93.3	320	2	Q8UDG0 AGRT5	Q8UDG0 agrobacteri
41	42	93.3	355	2	Q4IEH9 GIBZE	Q4IEH9 gibberella
42	42	93.3	359	2	Q7CXQ3 AGRT5	Q7CXQ3 agrobacteri
43	42	93.3	377	2	Q4SJB1 TETNG	Q4SJB1 tetraodon n
44	42	93.3	464	2	Q61C10 CAEBR	Q61C10 caenorhabdi
45	42	93.3	493	2	Q9UVD1 PNECA	Q9UVD1 pneumocysti
46	42	93.3	496	2	Q9RJX6 STRCO	Q9RJX6 streptomyce
47	42	93.3	542	2	Q6I587 ORYSA	Q6I587 oryza sativ
48	42	93.3	549	2	Q9FMF8 ARATH	Q9FMF8 arabidopsis
49	42	93.3	579	2	Q4RP94 TETNG	Q4RP94 tetraodon n
50	42	93.3	619	2	Q6AHW2 PNECA	Q6AHW2 pneumocysti
51	42	93.3	641	1	XKR6 HUMAN	Q5GH73 homo sapien
52	42	93.3	645	2	Q23374 ARATH	Q23374 arabidopsis
53	42	93.3	670	2	Q8RVE5 ORYSA	Q8RVE5 oryza sativ
54	42	93.3	673	2	Q5EAW6 XENLA	Q5EAW6 xenopus lae
55	42	93.3	736	1	DVL2 XENLA	P51142 xenopus lae
56	42	93.3	757	2	Q91YW2 MOUSE	Q91YW2 mus musculu
57	42	93.3	760	2	Q86A14 DICDI	Q86A14 dictyosteli
58	42	93.3	782	1	SOX30 MOUSE	Q8CGW4 mus musculu
59	42	93.3	782	2	Q5SSN4 MOUSE	Q5SSN4 mus musculu
60	42	93.3	858	2	Q4I5L0 GIBZE	Q4I5L0 gibberella
61	42	93.3	873	2	Q4IBP4 GIBZE	Q4IBP4 gibberella
62	42	93.3	902	2	Q6AHS3 PNECA	Q6AHS3 pneumocysti
63	42	93.3	902	2	Q6AHV9 PNECA	Q6AHV9 pneumocysti
64	42	93.3	933	2	Q9BL41 CABEL	Q9BL41 caenorhabdi
65	42	93.3	947	2	Q6AHS6 PNECA	Q6AHS6 pneumocysti
66	42	93.3	1158	2	Q4QGW6 LEIMA	Q4QGW6 leishmania
67	42	93.3	1264	2	Q6ZU65 HUMAN	Q6ZU65 homo sapien
68	42	93.3	1276	2	Q8MI49 FELCA	Q8MI49 felis silve
69	42	93.3	1276	2	Q8MKD8 FELCA	Q8MKD8 felis silve
70	42	93.3	1276	2	Q9N0Q0 FELCA	Q9N0Q0 felis silve
71	42	93.3	1277	1	NPC1_PIG	P56941 sus scrofa
72	42	93.3	1277	2	Q9GLC9 BOVIN	Q9GLC9 bos taurus
73	42	93.3	1277	2	Q7TMD4 MOUSE	Q7TMD4 mus musculu
74	42	93.3	1278	1	NPC1_HUMAN	O15118 homo sapien
75	42	93.3	1278	1	NPC1_MOUSE	O35604 mus musculu
76	42	93.3	1286	2	Q9TT75 RABIT	Q9TT75 oryctolagus
77	42	93.3	1289	2	Q59GR1 HUMAN	Q59GR1 homo sapien
78	42	93.3	1313	2	Q80WC1 MOUSE	Q80WC1 mus musculu
79	42	93.3	1393	2	Q8IPP9 DROME	Q8IPP9 drosophila
80	42	93.3	1417	2	Q7YU09 DROME	Q7YU09 drosophila
81	42	93.3	1461	2	Q4QC44 LEIMA	Q4QC44 leishmania
82	42	93.3	1621	2	Q9VIK7 DROME	Q9VIK7 drosophila
83	42	93.3	1745	1	ZO1_MOUSE	P39447 mus musculu
84	42	93.3	2691	2	Q618T4 CAEBR	Q618T4 caenorhabdi
85	42	93.3	2761	2	Q4P226 USTMA	Q4P226 ustilago ma
86	42	93.3	3467	2	Q8I218 PLAF7	Q8I218 plasmodium
87	42	93.3	4903	1	MLL3_MOUSE	Q8BRH4 mus musculu
88	41	91.1	33	2	O02832 9PRIM	O02832 gorilla gor
89	41	91.1	58	2	Q68GX9 CANLU	Q68GX9 canis lupus
90	41	91.1	58	2	Q68GY0 CANLA	Q68GY0 canis latra
91	41	91.1	70	2	Q6Z3I2 ORYSA	Q6Z3I2 oryza sativ
92	41	91.1	77	2	Q5QMJ7 ORYSA	Q5QMJ7 oryza sativ
93	41	91.1	95	1	BCD_DROSU	P55924 drosophila
94	41	91.1	101	2	Q6WZC9 BOVIN	Q6WZC9 bos taurus
95	41	91.1	104	2	Q5DM45 BOVIN	Q5DM45 bos taurus
96	41	91.1	107	2	O03900 9VIRU	O03900 bacterioph
97	41	91.1	127	2	Q69JN5 ORYSA	Q69JN5 oryza sativ
98	41	91.1	131	1	ASIP_CANFA	Q5UK76 canis fami1
99	41	91.1	131	1	ASIP_VULVU	P79407 vulpes vulp
100	41	91.1	133	2	Q5QLQ9 ORYSA	Q5QLQ9 oryza sativ
101	41	91.1	135	1	ASIP_FELCA	Q865F0 felis silve
102	41	91.1	135	2	Q6Z4C7 ORYSA	Q6Z4C7 oryza sativ
103	41	91.1	150	2	Q5IUH7_MAGGR	Q5IUH7 magnaporthe
104	41	91.1	150	2	Q8C005_MOUSE	Q8C005 mus musculu

105	41	91.1	159	2	Q9I024_PSEAE	Q9i024_pseudomonas	178	41	91.1	822	2	Q7XCT3_ORYSA	Q7xct3_oryza_sativ
106	41	91.1	176	2	Q8LM61_ORYSA	Q8lm61_oryza_sativ	179	41	91.1	822	2	Q94HZ3_ORYSA	Q94hz3_oryza_sativ
107	41	91.1	176	2	Q5ZE80_ORYSA	Q5ze80_oryza_sativ	180	41	91.1	834	2	Q55PX0_CRYNE	Q55px0_cryptococcu
108	41	91.1	181	2	Q7QYY0_GIALA	Q7qyy0_giardia_lam	181	41	91.1	834	2	Q5XDJ0_CRYNE	Q5xdj0_cryptococcu
109	41	91.1	187	2	Q654G3_ORYSA	Q654g3_oryza_sativ	182	41	91.1	837	1	FA40A_HUMAN	Q5vsl9_homo_sapien
110	41	91.1	190	2	Q7R5F8_GIALA	Q7r5f8_giardia_lam	183	41	91.1	837	1	FA40A_MACFA	Q9glz5_macaca_fasc
111	41	91.1	194	2	Q7R5F8_GIALA	Q7r5f8_giardia_lam	184	41	91.1	837	1	FA40A_MOUSE	Q8c079_mus_musculu
112	41	91.1	197	2	Q7XFS4_ORYSA	Q7xfs4_oryza_sativ	185	41	91.1	837	1	FA40A_PONPY	Q5r7s4_pongo_pygma
113	41	91.1	197	2	Q8S780_ORYSA	Q8s780_oryza_sativ	186	41	91.1	837	2	Q6NSB6_RHOPA	Q6n5b6_rhodopseu
114	41	91.1	205	2	Q4NRN0_9DELT	Q4nrn0_anaeromyxob	187	41	91.1	841	1	RELA_STRAT	Q85709_streptomyce
115	41	91.1	226	2	Q51TR6_MAGGR	Q51tr6_magnaporthe	188	41	91.1	851	2	Q55XY5_CRYNE	Q55xy5_cryptococcu
116	41	91.1	239	2	Q55C69_DICDI	Q55c69_dictyosteli	189	41	91.1	873	1	RX_DROME	Q9w2q1_drosophila
117	41	91.1	248	1	ICP34_HHV11	P36313_human herpes	190	41	91.1	873	2	Q5KM08_CRYNE	Q5km08_cryptococcu
118	41	91.1	248	1	RL1_HHV11	O12396_human herpes	191	41	91.1	874	2	Q5JLK9_ORYSA	Q5jlk9_oryza_sativ
119	41	91.1	251	2	Q52IU9_MAGGR	Q52iu9_magnaporthe	192	41	91.1	883	2	Q529K7_MAGGR	Q529k7_magnaporthe
120	41	91.1	255	2	Q90ZR5_BRARE	Q90zr5_brachydanio	193	41	91.1	895	2	Q7S902_NEUCR	Q7s902_neurospora
121	41	91.1	264	2	Q07393_MYCAV	Q07393_mycobacteri	194	41	91.1	896	2	Q66D18_HUMAN	Q96d18_homo_sapien
122	41	91.1	273	2	Q96KW5_HUMAN	Q96kw5_homo_sapien	195	41	91.1	902	2	Q68FL6_MOUSE	Q68fl6_mus_musculu
123	41	91.1	277	2	Q8WYH1_HUMAN	Q8wyh1_homo_sapien	196	41	91.1	972	2	Q9M2D4_ARATH	Q9m2d4_arabidopsis
124	41	91.1	279	2	Q6X235_9ALPH	Q6x235_bovine herp	197	41	91.1	979	1	RFX1_HUMAN	P22670_homo_sapien
125	41	91.1	288	2	Q5C0W6_SCHJA	Q5c0w6_schistosoma	198	41	91.1	993	2	Q7VVD6_BORPE	Q7vvd6_bordetella
126	41	91.1	292	2	Q5W6G5_ORYSA	Q5w6g5_oryza_sativ	199	41	91.1	1012	2	Q5VZ95_HUMAN	Q5vz95_homo_sapien
127	41	91.1	317	2	Q8H4Q2_ORYSA	Q8h4q2_oryza_sativ	200	41	91.1	1012	2	Q9UEF7_HUMAN	Q9uef7_homo_sapien
128	41	91.1	320	1	HXA4_HUMAN	Q00056_homo_sapien	201	41	91.1	1013	2	Q4SDR3_TETNG	Q4sdr3_tetraodon n
129	41	91.1	324	1	YBX2B_XENLA	P45441_xenopus lae	202	41	91.1	1014	2	Q8WP17_MACFA	Q8wp17_macaca_faac
130	41	91.1	327	2	Q8BHG7_MOUSE	Q8bhg7_m_mus muscu	203	41	91.1	1029	2	Q9YMX0_NPVL	Q9ymx0_lymantria d
131	41	91.1	338	2	Q6P0M8_HUMAN	Q6p0m8_homo_sapien	204	41	91.1	1125	2	Q4RJR1_TETNG	Q4rjr1_tetraodon n
132	41	91.1	342	2	Q4RE66_TETNG	Q4re66_tetraodon n	205	41	91.1	1165	2	Q86HN1_DICDI	Q86hn1_dictyosteli
133	41	91.1	343	2	Q6ZJE6_ORYSA	Q6zje6_oryza_sativ	206	41	91.1	1165	2	Q4QHC6_LEIMA	Q4qhc6_leishmania
134	41	91.1	353	2	Q6Q502_ORYSA	Q6q502_oryza_sativ	207	41	91.1	1188	2	Q7RW25_NEUCR	Q7rw25_neurospora
135	41	91.1	353	2	Q6Q7E3_ORYSA	Q6q7e3_oryza_sativ	208	41	91.1	1217	2	Q9ULL5_HUMAN	Q9ull5_homo_sapien
136	41	91.1	361	2	Q6ETV2_ORYSA	Q6ety2_oryza_sativ	209	41	91.1	1229	2	Q68H99_PIG	Q68h99_sus_scrofa
137	41	91.1	361	2	Q6IEP2_ORYSA	Q6iep2_oryza_sativ	210	41	91.1	1251	1	IRS1_CERAE	Q28224_cercopithec
138	41	91.1	380	2	Q7X9Y7_9CARY	Q7x9y7_fagopyrum s	211	41	91.1	1253	2	Q6DHV6_HUMAN	Q6dhv6_homo_sapien
139	41	91.1	381	2	Q7X9Y9_9CARY	Q7x9y9_fagopyrum g	212	41	91.1	1294	2	Q4SAM8_TETNG	Q4sam8_tetraodon n
140	41	91.1	381	2	Q7XXU0_ORYSA	Q7xku0_oryza_sativ	213	41	91.1	1338	1	IRS2_HUMAN	Q9y4h2_homo_sapien
141	41	91.1	389	2	Q59NU2_CANAL	Q59nu2_candida alb	214	41	91.1	1338	2	Q96RG4_HUMAN	Q96rg4_homo_sapien
142	41	91.1	390	2	Q59NP4_CANAL	Q59np4_candida alb	215	41	91.1	1339	2	Q96RG5_HUMAN	Q96rg5_homo_sapien
143	41	91.1	418	1	NCOA6_FAT	Q9jli4_r_nuclear r	216	41	91.1	1367	2	Q8WZV8_NEUCR	Q8wzv8_neurospora
144	41	91.1	427	1	BSH_DROME	Q04787_drosophila	217	41	91.1	1400	1	BRD4_MOUSE	Q9esu6_mus_musculu
145	41	91.1	429	2	Q4V6V3_DROME	Q4v6v3_drosophila	218	41	91.1	1420	2	Q9Y0Y8_DROME	Q9y0y8_drosophila
146	41	91.1	451	2	Q7Q6A6_ANOGA	Q7q6a6_anopheles g	219	41	91.1	1430	2	Q9H7P9_HUMAN	Q9h7p9_homo_sapien
147	41	91.1	456	2	Q7SH82_NEUCR	Q7sh82_neurospora	220	41	91.1	1506	2	Q8IRL0_DROME	Q8irl0_drosophila
148	41	91.1	458	2	Q4NSW5_9DELT	Q4nsw5_anaeromyxob	221	41	91.1	1641	2	Q5NCY0_MOUSE	Q5ncy0_mus_musculu
149	41	91.1	482	1	PDCD7_MOUSE	Q9wty1_mus_musculu	222	41	91.1	1641	2	Q4VC26_MOUSE	Q4vc26_mus_musculu
150	41	91.1	485	1	PDCD7_HUMAN	Q8n8d1_homo_sapien	223	41	91.1	1682	2	O15054_HUMAN	O15054_homo_sapien
151	41	91.1	485	2	Q6IEG3_HUMAN	Q6ieg3_homo_sapien	224	41	91.1	1998	1	MYST3_RAT	Q5tkr9_rattus norv
152	41	91.1	485	2	Q24I36_DROME	Q24i36_drosophila	225	41	91.1	2003	1	MYST3_MOUSE	Q8bz21_mus_musculu
153	41	91.1	485	2	Q9V552_DROME	Q9v552_drosophila	226	41	91.1	2004	1	MYST3_HUMAN	Q92794_h_histone a
154	41	91.1	502	2	Q8NFC6_HUMAN	Q8nfc6_homo_sapien	227	41	91.1	2133	2	Q8IQT4_DROME	Q8igt4_drosophila
155	41	91.1	520	2	Q4RQ72_TETNG	Q4rq72_tetraodon n	228	41	91.1	2314	2	Q4SJP0_TETNG	Q4sjp0_tetraodon n
156	41	91.1	551	2	Q8WPI8_MACFA	Q8wpi8_macaca fasc	229	41	91.1	2906	2	Q9WUH9_RAT	Q9wuh9_rattus norv
157	41	91.1	573	2	Q4S0L1_TETNG	Q4s0l1_tetraodon n	230	41	91.1	2907	1	FBN2_MOUSE	Q61555_mus_musculu
158	41	91.1	590	2	Q7Y2I4_ARATH	Q7y2i4_arabidopsis	231	41	91.1	2911	1	FBN2_HUMAN	P35556_homo_sapien
159	41	91.1	600	2	Q8N4J6_HUMAN	Q8n4j6_homo_sapien	232	41	91.1	3144	1	HD_HUMAN	P42858_homo_sapien
160	41	91.1	619	2	Q8UZE3_9GAMA	Q8uze3_cercopithec	233	41	91.1	3160	1	EP400_HUMAN	Q96l91_homo_sapien
161	41	91.1	621	1	Q7S8B6_NEUCR	Q7s8b6_neurospora	234	41	91.1	3469	2	Q9U4I2_DROME	Q9u4i2_drosophila
162	41	91.1	630	1	MECT1_MOUSE	Q6uuv7_mus_musculu	235	41	91.1	3571	2	Q4QI22_LEIMA	Q4qiz2_leishmania
163	41	91.1	634	1	MECT1_HUMAN	Q6uuv7_homo_sapien	236	41	91.1	3604	2	Q9VVK0_DROME	Q9vyk0_drosophila
164	41	91.1	641	2	Q5I2C0_MAGGR	Q5l2c0_magnaporthe	237	41	91.1	4911	1	MLL3_HUMAN	Q8nez4_homo_sapien
165	41	91.1	643	2	Q9LSI2_ARATH	Q9ls12_arabidopsis	238	40	88.9	43	2	P79301_PIG	P79301_sus_scrofa
166	41	91.1	645	2	Q5IY73_MAGGR	Q5iy73_magnaporthe	239	40	88.9	44	2	Q8HXS5_CAPCA	Q8hxs5_capreolus c
167	41	91.1	660	2	Q6IY16_CAEBR	Q6lj16_caenorhabdi	240	40	88.9	51	2	Q8HXS4_CAPCA	Q8hxs4_capreolus c
168	41	91.1	684	2	Q7YID3_ORYSA	Q7yid3_oryza_sativ	241	40	88.9	52	2	Q8HXT2_CAPCA	Q8hxt2_capreolus c
169	41	91.1	689	2	Q8WPL5_9UROC	Q8wpl5_oikopleura	242	40	88.9	52	2	Q8HXT4_CAPCA	Q8hxt4_capreolus c
170	41	91.1	708	2	Q5IYX7_MAGGR	Q5ily7_magnaporthe	243	40	88.9	67	2	Q6K657_ORYSA	Q6k657_oryza_sativ
171	41	91.1	711	2	Q7XT28_ORYSA	Q7xt28_oryza_sativ	244	40	88.9	70	2	Q9MAV4_ARATH	Q9mav4_arabidopsis
172	41	91.1	733	2	Q9VTX9_DROME	Q9vtx9_drosophila	245	40	88.9	89	2	Q6YTG3_ORYSA	Q6ytg3_oryza_sativ
173	41	91.1	736	1	CTCF_MOUSE	Q6l164_mus_musculu	246	40	88.9	105	2	Q7S3N9_NEUCR	Q7s3n9_neurospora
174	41	91.1	766	2	Q4RKI6_TETNG	Q4rki6_tetraodon n	247	40	88.9	109	2	Q40548_TOBAC	Q40548_nicotiana t
175	41	91.1	780	1	T22D2_HUMAN	Q75157_homo_sapien	248	40	88.9	110	2	Q7IDC1_DROYA	Q7idc1_drosophila
176	41	91.1	808	2	Q4P0G5_USTMA	Q4p0g5_ustilago ma	249	40	88.9	110	2	Q9W154_DROME	Q9w154_drosophila
177	41	91.1	814	2	Q4IR23_GIBZE	Q4ir23_gibberella	250	40	88.9	115	2	Q5AL62_CANAL	Q5al62_candida alb

251	40	88.9	116	2	Q53SV4_HUMAN	Q53ev4 homo sapien	324	40	88.9	226	2	Q61H53_CABBR	Q61h53 caenorhabdi
252	40	88.9	116	2	Q9VEI7_DROME	Q9vei7 drosophila	325	40	88.9	227	2	Q66L28_XENLA	Q66l28 xenopus lae
253	40	88.9	118	2	Q9H387_HUMAN	Q9h387 homo sapien	326	40	88.9	231	2	Q4Q5X2_LEIMA	Q4q5x2 leishmania
254	40	88.9	118	2	Q9D1Y5_MOUSE	Q9dly5 mus musculu	327	40	88.9	231	2	Q75IX1_ORYSA	Q75ix1 oryza sativ
255	40	88.9	121	2	Q4SP92_TETNG	Q4sp92 tetraodon n	328	40	88.9	235	2	Q8LLM9_ORYSA	Q8llm9 oryza sativ
256	40	88.9	122	2	Q6ZNY9_HUMAN	Q6zny9 homo sapien	329	40	88.9	239	2	Q51HW0_MAGGR	Q51hw0 magnaporthe
257	40	88.9	128	2	Q4IE09_GIBZE	Q4ie09 gibberella	330	40	88.9	240	2	P89268_GVXN	P89268 xestia c-ni
258	40	88.9	130	2	Q6ZSV2_HUMAN	Q6zsv2 homo sapien	331	40	88.9	242	2	Q43687_VIGUN	Q43687 vigna ungui
259	40	88.9	132	2	Q6ZNU7_HUMAN	Q6znu7 homo sapien	332	40	88.9	242	2	Q4RLI9_TETNG	Q4rl19 tetraodon n
260	40	88.9	135	2	Q95KE1_MACFA	Q95ke1 macaca fasc	333	40	88.9	243	2	Q84LD5_9ASTR	Q84ld5 chrysanthem
261	40	88.9	136	2	Q643A9_STRHY	Q643a9 streptomyce	334	40	88.9	246	2	Q6QE92_9MOLL	Q6qe92 pareledone
262	40	88.9	138	2	Q6ZUU5_HUMAN	Q6zuu5 homo sapien	335	40	88.9	247	2	Q75LM5_ORYSA	Q75lm5 oryza sativ
263	40	88.9	139	2	Q6Z571_ORYSA	Q6z571 oryza sativ	336	40	88.9	249	2	Q4US07_XANCP	Q4us07 xanthomonas
264	40	88.9	140	2	Q7XIT9_ORYSA	Q7xit9 oryza sativ	337	40	88.9	251	2	Q59RN8_CANAL	Q59rn8 candida alb
265	40	88.9	143	2	Q7PNB6_ANOGA	Q7pnb6 anopheles g	338	40	88.9	251	2	Q8L3S3_ORYSA	Q8l3s3 oryza sativ
266	40	88.9	147	2	Q5BWN9_SCHJA	Q5bwn9 schistosoma	339	40	88.9	252	2	Q5CD49_BRARE	Q5cd49 brachydanio
267	40	88.9	148	2	Q6ZUC5_HUMAN	Q6zuc5 homo sapien	340	40	88.9	254	2	Q61WP8_CABBR	Q61wp8 caenorhabdi
268	40	88.9	148	2	Q6ZWG0_HUMAN	Q6zwg0 homo sapien	341	40	88.9	259	2	Q9ARE2_FLABI	Q9are2 flaveria bi
269	40	88.9	151	2	Q5VQU5_ORYSA	Q5vqu5 oryza sativ	342	40	88.9	261	2	Q94LQ0_ORYSA	Q94lq0 oryza sativ
270	40	88.9	152	2	Q6ZUG4_HUMAN	Q6zug4 homo sapien	343	40	88.9	265	2	Q6UX52_HUMAN	Q6ux52 homo sapien
271	40	88.9	153	2	Q6ERN8_ORYSA	Q6ern8 oryza sativ	344	40	88.9	265	2	Q5C4Q7_SCHJA	Q5c4q7 schistosoma
272	40	88.9	155	2	Q9HBV3_HUMAN	Q9hbv3 homo sapien	345	40	88.9	266	2	Q7XL73_ORYSA	Q7xl73 oryza sativ
273	40	88.9	156	2	Q5T481_HUMAN	Q5t481 homo sapien	346	40	88.9	267	2	Q6DV99_CHICK	Q6dv99 gallus gall
274	40	88.9	158	2	Q9D VW0_9BACU	Q9dvw0 plutella xy	347	40	88.9	275	2	Q9W1U7_DROME	Q9w1u7 drosophila
275	40	88.9	161	2	Q75IJ3_ORYSA	Q75ij3 oryza sativ	348	40	88.9	275	2	Q7YU61_DROME	Q7yu61 drosophila
276	40	88.9	161	2	Q7XAL2_ORYSA	Q7xal2 oryza sativ	349	40	88.9	277	2	Q6K296_ORYSA	Q6k296 oryza sativ
277	40	88.9	162	2	Q9AWM3_ORYSA	Q9awm3 oryza sativ	350	40	88.9	280	2	Q86FT3_9TURB	Q86ft3 dugesia etr
278	40	88.9	165	2	Q77IW4_9NUCL	Q77iw4 helicoverpa	351	40	88.9	282	2	Q6WP44_MUSSP	Q6wp44 mus spretus
279	40	88.9	165	2	Q77LZ7_9NUCL	Q77lz7 helicoverpa	352	40	88.9	282	2	Q6WP45_MUSSP	Q6wp45 mus spretus
280	40	88.9	165	2	Q8V5R8_9NUCL	Q8v5r8 helicoverpa	353	40	88.9	282	2	Q6WP46_MUSSP	Q6wp46 mus spretus
281	40	88.9	165	2	Q91BW0_9NUCL	Q91bw0 helicoverpa	354	40	88.9	282	2	Q6WP48_MUSSP	Q6wp48 mus spretus
282	40	88.9	165	2	Q99GW5_9NUCL	Q99gw5 helicoverpa	355	40	88.9	283	2	Q5XEY8_ARATH	Q5xey8 arabidopsis
283	40	88.9	171	2	Q7DLN2_TOBAC	Q7dln2 nicotiana t	356	40	88.9	283	2	Q6Z1D9_ORYSA	Q6z1d9 oryza sativ
284	40	88.9	177	2	Q7XQ64_ORYSA	Q7xq64 oryza sativ	357	40	88.9	283	2	Q6Z6Q4_ORYSA	Q6z6q4 oryza sativ
285	40	88.9	177	2	Q8H062_ORYSA	Q8h062 oryza sativ	358	40	88.9	291	2	Q6BH87_DEBHA	Q6bh87 debaryomyce
286	40	88.9	178	2	Q69TQ6_ORYSA	Q69tq6 oryza sativ	359	40	88.9	296	2	Q7Q6Q5_ANOGA	Q7q6q5 anopheles g
287	40	88.9	180	2	Q8NAQ6_HUMAN	Q8nag6 homo sapien	360	40	88.9	297	2	Q9JLV7_MOUSE	Q9jlv7 mus musculu
288	40	88.9	182	2	Q7XXM6_ORYSA	Q7xxm6 oryza sativ	361	40	88.9	298	2	Q5QN41_ORYSA	Q5qn41 oryza sativ
289	40	88.9	184	2	Q5VRR2_ORYSA	Q5vrr2 oryza sativ	362	40	88.9	302	2	Q9C589_ARATH	Q9c589 arabidopsis
290	40	88.9	185	2	Q5QLQ3_ORYSA	Q5qlq3 oryza sativ	363	40	88.9	302	2	Q9C8B9_ARATH	Q9c8b9 arabidopsis
291	40	88.9	191	2	Q51ZW1_MAGGR	Q51zw1 magnaporthe	364	40	88.9	303	2	Q84711_9LUTE	Q84711 pea enation
292	40	88.9	192	2	Q7S944_NEUCR	Q7s944 neurospora	365	40	88.9	306	2	Q94HB0_ORYSA	Q94hb0 oryza sativ
293	40	88.9	196	2	Q08195_TOBAC	Q08195 nicotiana t	366	40	88.9	310	2	Q527E0_MAGGR	Q527e0 magnaporthe
294	40	88.9	196	2	Q84UR6_ORYSA	Q84ur6 oryza sativ	367	40	88.9	310	2	Q5ZDE8_ORYSA	Q5zde8 oryza sativ
295	40	88.9	197	2	Q4I3A7_GIBZE	Q4i3a7 gibberella	368	40	88.9	312	2	Q4NL12_9MICC	Q4nl12 arthrobacte
296	40	88.9	198	2	Q4I7M8_GIBZE	Q4i7m8 gibberella	369	40	88.9	313	1	CDX2_HUMAN	Q99626 homo sapien
297	40	88.9	203	2	Q4QBL7_LEIMA	Q4qbl7 leishmania	370	40	88.9	313	1	ENV1_MOUSE	P10404 mus musculu
298	40	88.9	205	2	Q69W83_ORYSA	Q69w83 oryza sativ	371	40	88.9	313	2	Q5VTU7_HUMAN	Q5vtu7 homo sapien
299	40	88.9	206	2	Q825Z2_STRAW	Q825z2 streptomyce	372	40	88.9	315	2	Q5ISK4_MACFA	Q5isk4 macaca fasc
300	40	88.9	207	2	Q8EFY6_SHEON	Q8efy6 shewanella	373	40	88.9	316	2	Q57XT4_9TRYP	Q57xt4 trypanosoma
301	40	88.9	209	2	Q08194_TOBAC	Q08194 nicotiana t	374	40	88.9	317	2	Q6PB42_MOUSE	Q6pb42 mus musculu
302	40	88.9	212	2	Q7Y1G1_ORYSA	Q7y1g1 oryza sativ	375	40	88.9	319	2	Q9FLU5_ARATH	Q9flu5 arabidopsis
303	40	88.9	213	2	Q6SQE4_HDV	Q6sqe4 hepatitis d	376	40	88.9	320	2	Q6ZSA5_HUMAN	Q6zsa5 homo sapien
304	40	88.9	213	2	Q760D0_HDV	Q760d0 hepatitis d	377	40	88.9	320	2	Q8L9K2_ARATH	Q8l9k2 arabidopsis
305	40	88.9	213	2	Q760D2_HDV	Q760d2 hepatitis d	378	40	88.9	322	2	Q4NU48_9DELT	Q4nu48 anaeromyxob
306	40	88.9	213	2	Q760D3_HDV	Q760d3 hepatitis d	379	40	88.9	324	2	Q8LRH2_ORYSA	Q8lrh2 oryza sativ
307	40	88.9	213	2	Q760D4_HDV	Q760d4 hepatitis d	380	40	88.9	325	2	Q9BSN0_HUMAN	Q9bsn0 homo sapien
308	40	88.9	213	2	Q760D5_HDV	Q760d5 hepatitis d	381	40	88.9	325	2	Q4RTH8_TETNG	Q4rth8 tetraodon n
309	40	88.9	213	2	Q9PWX0_HDV	Q9pwx0 hepatitis d	382	40	88.9	326	2	Q91WA8_MOUSE	Q91wa8 mus musculu
310	40	88.9	213	2	Q9PX07_HDV	Q9px07 hepatitis d	383	40	88.9	328	2	Q5VVK1_ORYSA	Q5vnk1 oryza sativ
311	40	88.9	213	2	Q9PZV7_HDV	Q9pzv7 hepatitis d	384	40	88.9	328	2	Q64262_MOUSE	Q64262 mus musculu
312	40	88.9	213	2	Q9PZV8_HDV	Q9pzv8 hepatitis d	385	40	88.9	329	2	Q8H8U0_ORYSA	Q8h8u0 oryza sativ
313	40	88.9	213	2	Q9PZV9_HDV	Q9pzv9 hepatitis d	386	40	88.9	334	2	Q7TZB1_MYCBO	Q7tzb1 mycobacteri
314	40	88.9	213	2	Q9PZW0_HDV	Q9pzw0 hepatitis d	387	40	88.9	334	2	O07721_MYCTU	O07721 mycobacteri
315	40	88.9	213	2	Q9PZW1_HDV	Q9pzw1 hepatitis d	388	40	88.9	335	2	Q6H830_ORYSA	Q6h830 oryza sativ
316	40	88.9	213	2	Q9PZW2_HDV	Q9pzw2 hepatitis d	389	40	88.9	340	2	Q922N8_MOUSE	Q922n8 mus musculu
317	40	88.9	213	2	Q9PZW3_HDV	Q9pzw3 hepatitis d	390	40	88.9	340	2	Q6KAU4_MOUSE	Q6kau4 mus musculu
318	40	88.9	215	2	Q61685_MOUSE	Q61685 mus musculu	391	40	88.9	342	2	Q5VPC0_ORYSA	Q5vpc0 oryza sativ
319	40	88.9	217	2	Q7XU26_ORYSA	Q7xu26 oryza sativ	392	40	88.9	343	2	Q8LN10_ORYSA	Q8lni0 oryza sativ
320	40	88.9	221	2	Q8N6S7_HUMAN	Q8n6s7 homo sapien	393	40	88.9	344	1	TILS_SYNEL	Q8dlh2 synechococc
321	40	88.9	221	2	O49946_SOLTU	O49946 solanum tub	394	40	88.9	344	2	Q9H7P6_HUMAN	Q9h7p6 homo sapien
322	40	88.9	222	2	Q95S23_DROME	Q95s23 drosophila	395	40	88.9	345	2	Q7Q4G5_ANOGA	Q7q4g5 anopheles g
323	40	88.9	222	2	Q4TA38_TETNG	Q4ta38 tetraodon n	396	40	88.9	346	1	E2F5_HUMAN	Q15329 homo sapien

324	40	88.9	226	2	Q61H53_CABBR	Q61h53 caenorhabdi
325	40	88.9	227	2	Q66L28_XENLA	Q66l28 xenopus lae
326	40	88.9	231	2	Q4Q5X2_LEIMA	Q4q5x2 leishmania
327	40	88.9	231	2	Q75IX1_ORYSA	Q75ix1 oryza sativ
328	40	88.9	235	2	Q8LLM9_ORYSA	Q8llm9 oryza sativ
329	40	88.9	239	2	Q51HW0_MAGGR	Q51hw0 magnaporthe
330	40	88.9	240	2	P89268_GVXN	P89268 xestia c-ni
331	40	88.9	242	2	Q43687_VIGUN	Q43687 vigna ungui
332	40	88.9	242	2	Q4RLI9_TETNG	Q4rl19 tetraodon n
333	40	88.9	243	2	Q84LD5_9ASTR	Q84ld5 chrysanthem
334	40	88.9	246	2	Q6QE92_9MOLL	Q6qe92 pareledone
335	40	88.9	247	2	Q75LM5_ORYSA	Q75lm5 oryza sativ
336	40	88.9	249	2	Q4US07_XANCP	Q4us07 xanthomonas
337	40	88.9	251	2	Q59RN8_CANAL	Q59rn8 candida alb
338	40	88.9	251	2	Q8L3S3_ORYSA	Q8l3s3 oryza sativ
339	40	88.9	252	2	Q5CD49_BRARE	Q5cd49 brachydanio
340	40	88.9	254	2	Q61WP8_CABBR	Q61wp8 caenorhabdi
341	40	88.9	259	2	Q9ARE2_FLABI	Q9are2 flaveria bi
342	40	88.9	261	2	Q94LQ0_ORYSA	Q94lq0 oryza sativ
343	40	88.9	265	2	Q6UX52_HUMAN	Q6ux52 homo sapien
344	40	88.9	265	2	Q5C4Q7_SCHJA	Q5c4q7 schistosoma
345	40	88.9	266	2	Q7XL73_ORYSA	Q7xl73 oryza sativ
346	40	88.9	267	2	Q6DV99_CHICK	Q6dv99 gallus gall
347	40	88.9	275	2	Q9W1U7_DROME	Q9w1u7 drosophila
348	40	88.9	275	2	Q7YU61_DROME	Q7yu61 drosophila
349	40	88.9	277	2	Q6K296_ORYSA	Q6k296 oryza sativ
350	40	88.9	280	2	Q86FT3_9TURB	Q86ft3 dugesia etr
351	40	88.9	282	2	Q6WP44_MUSSP	Q6wp44 mus spretus
352	40	88.9	282	2	Q6WP45_MUSSP	Q6wp45 mus spretus
353	40	88.9	282	2	Q6WP46_MUSSP	Q6wp46 mus spretus
354	40	88.9	282	2	Q6WP48_MUSSP	Q6wp48 mus spretus
355	40	88.9	283	2	Q5XEY8_ARATH	Q5xey8 arabidopsis
356	40	88.9	283	2	Q6Z1D9_ORYSA	Q6z1d9 oryza sativ
357	40	88.9	283	2	Q6Z6Q4_ORYSA	Q6z6q4 oryza sativ
358	40	88.9	291	2	Q6BH87_DEBHA	Q6bh87 deb

397	40	88.9	349	2	Q9P555_NEUCR	Q9p555 neurospora	470	40	88.9	444	2	Q5AGS1_CANAL	Q5ags1 candida alb
398	40	88.9	350	1	HNF316_HUMAN	P55318 homo sapien	471	40	88.9	444	2	Q5MNP0_9GAMR	Q5mnp0 murine leuk
399	40	88.9	350	1	Q53F16_HUMAN	Q53f16 homo sapien	472	40	88.9	445	2	Q5MNN8_9GAMR	Q5mn8 murine leuk
400	40	88.9	351	1	KLf2_RAT	Q9et58 rattus norv	473	40	88.9	451	2	Q5MNN8_9GAMR	Q5mn8 murine leuk
401	40	88.9	352	2	Q6ISE7_HUMAN	Q6ise7 homo sapien	474	40	88.9	451	2	Q4U2V9_CHLRE	Q4u2v9 chlamydomon
402	40	88.9	353	2	Q6TRW9_CHICK	Q6trw9 gallus gall	475	40	88.9	453	2	Q96JE9_HUMAN	Q96je9 homo sapien
403	40	88.9	354	1	ATOH1_HUMAN	Q92858 homo sapien	476	40	88.9	454	2	Q69022_9GAMA	Q69022 human herpe
404	40	88.9	354	1	KLf2_MOUSE	Q60843 mus musculu	477	40	88.9	455	2	Q90120_9GAMR	Q90120 mink cell f
405	40	88.9	354	2	Q55FU3_DICDI	Q55fu3 dictyosteli	478	40	88.9	456	2	Q8IRW4_DROME	Q8irw4 drosophila
406	40	88.9	355	2	Q4IJ14_GIBZE	Q4ij14 gibberella	479	40	88.9	457	1	CHKA_HUMAN	P35790 homo sapien
407	40	88.9	355	2	Q4T8C9_TETNG	Q4t8c9 tetraodon n	480	40	88.9	457	2	Q4PIE9_USTMA	Q4pie9 ustilago ma
408	40	88.9	356	1	ATOH1_PANTR	Q5is79 pan troglod	481	40	88.9	463	2	Q6DHH6_BRARE	Q6dhh6 brachydanio
409	40	88.9	356	1	ENV_FRSPB	P31793 friend sple	482	40	88.9	469	1	FOXGA_HUMAN	P55316 homo sapien
410	40	88.9	356	2	Q6TRX0_CHICK	Q6trx0 gallus gall	483	40	88.9	474	2	Q35102_RATRT	Q35102 rattus ratt
411	40	88.9	356	2	Q87044_9GAMR	Q87044 spleen focu	484	40	88.9	477	1	FOXGB_HUMAN	P55315 homo sapien
412	40	88.9	360	2	Q4QDS7_LEIMA	Q4qds7 leishmania	485	40	88.9	477	2	Q851M1_ORYSA	Q851m1 oryza sativ
413	40	88.9	361	2	Q93VV8_ORYSA	Q93vv8 oryza sativ	486	40	88.9	477	2	Q96800_FRMCV	Q96800 friend mink
414	40	88.9	369	2	Q61Z90_CAEBR	Q61z90 caenorhabdi	487	40	88.9	480	1	FOXGB_MOUSE	Q00939 rattus norv
415	40	88.9	372	2	Q60TD1_CAEBR	Q60td1 caenorhabdi	488	40	88.9	481	1	FOXGB_MOUSE	Q60987 mus musculu
416	40	88.9	374	2	Q525P5_MAGGR	Q525p5 magnaporthe	489	40	88.9	485	2	Q7XEI9_ORYSA	Q7xei9 oryza sativ
417	40	88.9	374	2	QSTU44_ANOGA	Qstu44 anopheles g	490	40	88.9	486	2	Q8LAL7_ARATH	Q8lal7 arabidopsis
418	40	88.9	376	2	Q4XOM9_ASPPU	Q4xom9 aspergillus	491	40	88.9	486	2	Q9FNN5_ARATH	Q9fnn5 arabidopsis
419	40	88.9	376	2	Q4SD35_TETNG	Q4sd35 tetraodon n	492	40	88.9	489	2	Q86XT7_HUMAN	Q86xt7 homo sapien
420	40	88.9	377	2	Q6GUH1_HEVBR	Q6guh1 hevea bras	493	40	88.9	489	2	Q9XUM8_CAEBL	Q9xum8 caenorhabdi
421	40	88.9	378	2	Q6T9Z5_CHICK	Q6t9z5 gallus gall	494	40	88.9	490	2	Q8K557_MOUSE	Q8k557 mus musculu
422	40	88.9	379	2	Q504Y4_HUMAN	Q504y4 homo sapien	495	40	88.9	490	2	Q6PDH9_MOUSE	Q6pdh9 mus musculu
423	40	88.9	387	2	Q82LR9_STRAW	Q82lr9 streptomyce	496	40	88.9	491	2	Q69JG3_ORYSA	Q69jg3 oryza sativ
424	40	88.9	388	2	Q8MSD8_DROME	Q8msd8 drosophila	497	40	88.9	493	2	Q5BHV6_DROME	Q5bhv6 drosophila
425	40	88.9	388	2	Q84T11_ORYSA	Q84t11 oryza sativ	498	40	88.9	493	2	Q09708_PEAHV	Q09708 pea enation
426	40	88.9	389	2	Q73TN8_MYCPA	Q73tn8 mycobacteri	499	40	88.9	493	2	Q93184_PEAHV	Q93184 pea enation
427	40	88.9	390	1	DAF1_MOUSE	Q61475 mus musculu	500	40	88.9	495	1	PO3F3_MOUSE	P31361 mus musculu
428	40	88.9	390	1	OST3B_HUMAN	Q9Y662 homo sapien	501	40	88.9	497	2	Q5QEU3_9ACTO	Q5qeu3 streptomyce
429	40	88.9	390	1	OST3B_MOUSE	Q9qz86 mus musculu	502	40	88.9	498	2	Q68SS1_9AGAR	Q68ss1 pleurotus d
430	40	88.9	390	2	Q5NCRO_MOUSE	Q5ncr0 mus musculu	503	40	88.9	500	1	PO3F3_HUMAN	P20264 homo sapien
431	40	88.9	390	2	Q921P0_MOUSE	Q921p0 mus musculu	504	40	88.9	500	2	Q4ZG25_HUMAN	Q4zg25 homo sapien
432	40	88.9	390	2	Q4FJ84_MOUSE	Q4fj84 mus musculu	505	40	88.9	500	2	Q5QLD5_ORYSA	Q5qld5 oryza sativ
433	40	88.9	391	2	Q54M74_DICDI	Q54m74 dictyosteli	506	40	88.9	501	2	Q4RWZ6_TETNG	Q4rwz6 tetraodon n
434	40	88.9	394	2	Q8ZZ49_PYRAE	Q8zz49 pyrobaculum	507	40	88.9	503	2	Q8WSU5_9METZ	Q8wsu5 ephydatia f
435	40	88.9	396	2	Q522C4_MAGGR	Q522c4 magnaporthe	508	40	88.9	504	2	Q8H2T7_ORYSA	Q8h2t7 oryza sativ
436	40	88.9	397	2	Q85733_FRSPV	Q85733 friend sple	509	40	88.9	505	2	Q4P594_USTMA	Q4p594 ustilago ma
437	40	88.9	402	2	Q9UPV1_HUMAN	Q9upv1 homo sapien	510	40	88.9	505	2	WRK10_ARATH	Q4p584 arabidopsis
438	40	88.9	403	2	Q5JN13_ORYSA	Q5jnl3 oryza sativ	511	40	88.9	506	1	Q5MNNK2_9GAMR	Q5mnk2 murine leuk
439	40	88.9	406	2	Q6H4X7_ORYSA	Q6h4x7 oryza sativ	512	40	88.9	509	2	Q5VR46_ORYSA	Q5vr46 oryza sativ
440	40	88.9	408	1	ENV_RSFFV	P03389 rauscher sp	513	40	88.9	520	2	Q5MNNL0_9GAMR	Q5mn10 murine leuk
441	40	88.9	408	1	INHEB_BOVIN	P42917 bos taurus	514	40	88.9	526	2	Q7SEK3_NEUCR	Q7sek3 neurospora
442	40	88.9	410	2	Q61K70_CAEBR	Q61k70 caenorhabdi	515	40	88.9	527	2	Q54ND4_DICDI	Q54nd4 dictyosteli
443	40	88.9	410	2	Q7PXR9_ANOGA	Q7pxr9 anopheles g	516	40	88.9	532	2	Q6YWA3_ORYSA	Q6ywa3 oryza sativ
444	40	88.9	410	2	Q92791_FRSPV	Q92791 friend sple	517	40	88.9	533	1	CDPK2_ORYSA	P53683 oryza sativ
445	40	88.9	411	2	Q60R98_CAEBR	Q60r98 caenorhabdi	518	40	88.9	533	2	Q8GTY8_ORYSA	Q8gt8 oryza sativ
446	40	88.9	412	1	ALF_PETHY	O22621 petunia hyb	519	40	88.9	534	1	FTZF1_BOMMO	P49867 bombyx mori
447	40	88.9	412	2	Q651H6_ORYSA	Q651h6 oryza sativ	520	40	88.9	535	2	Q5MNNL0_9GAMR	Q5mn10 murine leuk
448	40	88.9	413	2	Q9SVU6_ARATH	Q9svu6 arabidopsis	521	40	88.9	535	2	Q5MNNL8_9GAMR	Q5mn18 murine leuk
449	40	88.9	415	2	Q9VT88_DROME	Q9vt88 drosophila	522	40	88.9	535	2	Q5MNNM2_9GAMR	Q5mn2 murine leuk
450	40	88.9	416	1	HXD3_HUMAN	P31249 homo sapien	523	40	88.9	535	2	Q5MNN2_9GAMR	Q5mn2 murine leuk
451	40	88.9	416	2	Q95TV8_DROME	Q95tv8 drosophila	524	40	88.9	535	2	Q5MNN4_9GAMR	Q5mn4 murine leuk
452	40	88.9	417	1	HXD3_MOUSE	P09027 mus musculu	525	40	88.9	537	2	Q55G00_DICDI	Q55g00 dictyosteli
453	40	88.9	417	2	Q8SXI7_DROME	Q8sxi7 drosophila	526	40	88.9	537	2	Q4NTG0_9DELT	Q4ntg0 anaeromyxob
454	40	88.9	422	2	Q9HXT5_PSEAE	Q9hxt5 pseudomonas	527	40	88.9	546	2	Q86XP8_HUMAN	Q86xp8 homo sapien
455	40	88.9	425	2	Q6YX84_ORYSA	Q6yx84 oryza sativ	528	40	88.9	546	2	Q40451_TOBAC	Q40451 nicotiana t
456	40	88.9	425	2	Q8RUJ8_ORYSA	Q8ruj8 oryza sativ	529	40	88.9	549	2	O61707_CAEBL	O61707 caenorhabdi
457	40	88.9	427	2	Q64ON2_MOUSE	Q64on2 mus musculu	530	40	88.9	553	2	Q9W153_DROME	Q9w153 drosophila
458	40	88.9	428	1	FOXB2_MOUSE	Q64733 mus musculu	531	40	88.9	553	2	Q93P58_MYXXA	Q93p58 myxococcus
459	40	88.9	431	1	KNAT3_ARATH	P48000 arabidopsis	532	40	88.9	556	2	Q6BIQ2_DEBHA	Q6biq2 debaryomyce
460	40	88.9	431	2	Q8LDC1_ARATH	Q8ldc1 arabidopsis	533	40	88.9	557	1	WRT4_CAEBL	Q94129 caenorhabdi
461	40	88.9	432	2	Q96HM7_HUMAN	Q96hm7 homo sapien	534	40	88.9	558	2	Q96D88_HUMAN	Q96d88 homo sapien
462	40	88.9	432	2	Q5VVV0_HUMAN	Q5vvv0 homo sapien	535	40	88.9	559	2	Q943J0_ORYSA	Q943j0 oryza sativ
463	40	88.9	435	2	Q5MNNK0_9GAMR	Q5mnk0 murine leuk	536	40	88.9	561	2	Q9H2M8_HUMAN	Q9h2m8 homo sapien
464	40	88.9	436	2	Q7SBR4_NEUCR	Q7sbr4 neurospora	537	40	88.9	562	2	Q8NAP0_HUMAN	Q8naf0 homo sapien
465	40	88.9	439	2	Q6X2S6_RETFL	Q6x2s6 reticuliter	538	40	88.9	562	2	Q54EY4_DICDI	Q54ey4 dictyosteli
466	40	88.9	439	2	Q6YZ63_ORYSA	Q6yz63 oryza sativ	539	40	88.9	562	2	Q9L6Y0_MYXXA	Q9l6y0 myxococcus
467	40	88.9	440	2	O01672_ASCSU	O01672 ascaris suu	540	40	88.9	562	2	Q80VM4_MOUSE	Q80vm4 mus musculu
468	40	88.9	440	2	Q85561_9GAMR	Q85561 mink cell f	541	40	88.9	564	2	Q7KZY0_HUMAN	Q7kzy0 homo sapien
469	40	88.9	442	2	Q8N7F2_HUMAN	Q8n7f2 homo sapien	542	40	88.9	568	2	Q55J33_CRYNE	Q55j33 cryptococcu

470	40	88.9	444	2	Q5AGS1_CANAL	Q5ags1 candida alb
471	40	88.9	444	2	Q5MNP0_9GAMR	Q5mnp0 murine leuk
472	40	88.9	445	2	Q5MNN8_9GAMR	Q5mn8 murine leuk
473	40	88.9	451	2	Q969F2_HUMAN	Q969f2 homo sapien
474	40	88.9	451	2	Q4U2V9_CHLRE	Q4u2v9 chlamydomon
475	40	88.9	453	2	Q96JE9_HUMAN	Q96je9 homo sapien
476	40	88.9	454	2	Q69022_9GAMA	Q69022 human herpe
477	40	88.9	455	2	Q90120_9GAMR	Q90120 mink cell f
478	40	88.9	456	2	Q8IRW4_DROME	Q8irw4 drosophila
479	40	88.9	457	1	CHKA_HUMAN	P35790 homo sapien
480	40	88.9	457	2	Q4PIE9_USTMA	Q4pie9 ustilago ma
481	40	88.9	463	2	Q6DHH6_BRARE	Q6dhh6 brachydanlo
482	40	88.9	469	1	FOXGA_HUMAN	P55316 homo sapien
483	40	88.9	474	2	Q35102_RATRT	Q35102 rattus ratt
484	40	88.9	477	1	FOXGB_HUMAN	P55315 homo sapien
485	40	88.9	477	2	Q851M1_ORYSA	Q851m1 oryza sativ
486	40	88.9	477	2	Q96800_FRMCV	Q96800 friend mink
487	40	88.9	480	1	FOXGB_RAT	Q00939 rattus norv
488	40	88.9	481	1	FOXGB_MOUSE	Q60987 mus musculu
489	40	88.9	485	2	Q7XEI9_ORYSA	Q7xei9 oryza sativ
490	40	88.9	486	2	Q8LAL7_ARATH	Q8lal7 arabidopsis
491	40	88.9	486	2	Q9FNN5_ARATH	Q9fnn5 arabidopsis
492	40	88.9	489	2	Q86XT7_HUMAN	Q86xt7 homo sapien
493	40	88.9	489	2	Q9XUM8_CAEBL	Q9xum8 caenorhabdi
494	40	88.9	490	2	Q8K557_MOUSE	Q8k557 mus musculu
495	40	88.9	490	2	Q6PDH9_MOUSE	Q6pdh9 mus musculu
496	40	88.9	491	2	Q69JG3_ORYSA	Q69jg3 oryza sativ
497	40	88.9	493	2	Q5BHV6_DROME	Q5bhv6 drosophila
498	40	88.9	493	2	Q09708_PEAHV	Q09708 pea enation
499	40	88.9	493	2	Q93184_PEAHV	Q93184 pea enation
500	40	88.9	495	1	PO3F3_MOUSE	P31361 mus musculu
501	40	88.9	497	2	Q5QEU3_9ACTO	Q5geu3 streptomyce
502	40	88.9	498	2	Q68SS1_9AGAR	Q68ss1 pleurotus d
503	40	88.9	500	1	PO3F3_HUMAN	P20264 homo sapien
504	40	88.9	500	2	Q4ZG25_HUMAN	Q4zg25 homo sapien
505	40	88.9	500	2	Q5QLD5_ORYSA	Q5qld5 oryza sativ
506	40	88.9	501	2	Q4RWZ6_TETNG	Q4rwz6 tetraodon n
507	40	88.9	503	2	Q8WSU5_9METZ	Q8wsu5 ephydatia f
508	40	88.9	504	2	Q8H2T7_ORYSA	Q8h2t7 oryza sativ
509	40	88.9	505	2	Q4PSS4_USTMA	Q4pss4 ustilago ma
510	40	88.9	506	1	WRK10_ARATH	Q91g05 arabidopsis
511	40	88.9	509	2	Q5MNK2_9GAMR	Q5mnk2 murine leuk
512	40	88.9	520	2	Q5VVR46_ORYSA	Q5vr46 oryza sativ
513	40	88.9	526	2	Q7SEK3_NEUCR	Q7sek3 neurospora
514	40	88.9	527	2	Q54ND4_DICDI	Q54nd4 dictyostell
515	40	88.9	529	2	Q5MNM0_9GAMR	Q5mnm0 murine leuk
516	40	88.9	532	2	Q6YWA3_ORYSA	Q6ywa3 oryza sativ
517	40	88.9	533	1	CDPK2_ORYSA	P53683 oryza sativ
518	40	88.9	533	2	Q8GTY8_ORYSA	Q8gty8 oryza sativ
519	40	88.9	534	1	FTZP1_BOMMO	P49867 bombyx mori
520	40	88.9	535	2	Q5MNL0_9GAMR	Q5mn10 murine leuk
521	40	88.9	535	2	Q5MNL8_9GAMR	Q5mn18 murine leuk
522	40	88.9	535	2	Q5MNM2_9GAMR	Q5mn12 murine leuk
523	40	88.9	535	2	Q5MNN2_9GAMR	Q5mn22 murine leuk
524	40	88.9	535	2	Q5MNN4_9GAMR	Q5mn24 murine leuk
525	40	88.9	537	2	Q55G00_DICDI	Q55g00 dictyostell
526	40	88.9	537	2	Q4NTG0_9DELT	Q4ntg0 anaeromyxob
527	40	88.9	546	2	Q86XP8_HUMAN	Q86xp8 homo sapien
528	40	88.9	546	2	Q40451_TOBAC	Q40451 nicotiana t
529	40	88.9	549	2	Q61707_CABEL	O61707 caenorhabdi
530	40	88.9	553	2	Q9W153_DROME	Q9w153 drosophila
531	40	88.9	553	2	Q93P58_MYXXA	Q93p58 myxococcus
532	40	88.9	556	2	Q6BIQ2_DEBHA	Q6biq2 debaryomyce
533	40	88.9	557	1	WRT4_CAEBL	Q94129 caenorhabdi
534	40	88.9	558	2	Q96D88_HUMAN	Q96d88 homo sapien
535	40	88.9	559	2	Q943J0_ORYSA	Q943j0 oryza sativ
536	40	88.9	561	2	Q9H2M8_HUMAN	Q9h2m8 homo sapien
537	40	88.9	562	2	Q8NAF0_HUMAN	Q8naf0 homo sapien
538	40	88.9	562	2	Q54EY4_DICDI	Q54ey4 dictyostell
539	40	88.9	562	2	Q9L6Y0_MVXXA	Q9l6y0 myxococcus
540	40	88.9	562	2	Q80VM4_MOUSE	Q80vm4 mus musculu
541	40	88.9	564	2	Q7KZY0_HUMAN	Q7kzy0 homo sapien
542	40	88.9	568	2	Q55J33_CRYNE	Q55j33 cryptococcu

543	40	88.9	568	2	Q5KCM3_CRYNE	Q5kcm3	cryptococcu
544	40	88.9	572	2	Q8GZY8_ORYSA	Q8gzy8	oryza sativ
545	40	88.9	576	2	Q5ZK52_CHICK	Q5zk52	gallus gall
546	40	88.9	579	1	SOCS7_MOUSE	Q8vhq2	mus musculus
547	40	88.9	582	2	Q8WSJ3_MANSE	Q8wsj3	manduca sex
548	40	88.9	582	2	Q63ZX2_MOUSE	Q63zx2	mus musculus
549	40	88.9	585	2	Q8WTX2_HUMAN	Q8wtx2	homo sapien
550	40	88.9	585	2	Q9BWN1_HUMAN	Q9bwn1	homo sapien
551	40	88.9	586	2	Q94GR7_ORYSA	Q94gr7	oryza sativ
552	40	88.9	586	2	Q58EN3_BRARE	Q58en3	brachydanio
553	40	88.9	588	2	Q5AM01_CANAL	Q5am01	candida alb
554	40	88.9	588	2	Q948Y8_VOLCA	Q948y8	volvox cart
555	40	88.9	588	2	Q9SNE9_ARATH	Q9sne9	arabidopsis
556	40	88.9	589	2	Q5ALK4_CANAL	Q5alk4	candida alb
557	40	88.9	590	2	Q61876_MOUSE	Q61876	mus musculus
558	40	88.9	593	1	ALU6_HUMAN	P39193	homo sapien
559	40	88.9	594	2	Q9VEP4_DROME	Q9vep4	drosophila
560	40	88.9	597	2	Q55VE2_CRYNE	Q55ve2	cryptococcu
561	40	88.9	597	2	Q5KKU4_CRYNE	Q5kku4	cryptococcu
562	40	88.9	600	2	Q7YXU4_DICDI	Q7yxu4	dictyosteli
563	40	88.9	601	2	Q7S9Q4_NEUCR	Q7s9q4	neurospora
564	40	88.9	602	2	Q4G3H3_BRARE	Q4g3h3	brachydanio
565	40	88.9	604	2	Q5R6Z2_PONPY	Q5r6z2	pongo pygma
566	40	88.9	604	2	Q651Z0_ORYSA	Q651z0	oryza sativ
567	40	88.9	605	1	NCKX4_HUMAN	Q8nff2	homo sapien
568	40	88.9	605	1	NCKX4_MOUSE	Q8cgg8	mus musculus
569	40	88.9	606	2	Q852P0_VOLCA	Q852p0	volvox cart
570	40	88.9	609	2	Q4RH16_TETNG	Q4rh16	tetraodon n
571	40	88.9	610	2	Q7QA42_ANOGA	Q7qa42	anopheles g
572	40	88.9	610	2	Q66650_9GAMA	Q66650	equid herpe
573	40	88.9	612	2	Q7TPN9_MOUSE	Q7tpn9	mus musculus
574	40	88.9	613	2	Q6Z697_ORYSA	Q6z697	oryza sativ
575	40	88.9	615	2	Q9UDW8_HUMAN	Q9udw8	homo sapien
576	40	88.9	615	2	Q7PIF2_ANOGA	Q7pif2	anopheles g
577	40	88.9	616	1	VGf_HUMAN	O15240	homo sapien
578	40	88.9	616	2	Q8T0J3_DROME	Q8t0j3	drosophila
579	40	88.9	617	1	VGf_RAT	P20156	rattus norv
580	40	88.9	619	2	Q6H672_ORYSA	Q6h672	oryza sativ
581	40	88.9	622	1	LAM0_DROME	P08928	drosophila
582	40	88.9	622	2	Q9VMQ0_DROME	Q9vmq0	drosophila
583	40	88.9	625	2	Q52F37_MAGGR	Q52f37	magnaporth
584	40	88.9	625	2	Q7QC98_ANOGA	Q7qc98	anopheles g
585	40	88.9	628	2	Q5SHU0_THET8	Q5shu0	thermus the
586	40	88.9	630	2	Q59VE3_CANAL	Q59ve3	candida alb
587	40	88.9	630	2	Q59V98_CANAL	Q59v98	candida alb
588	40	88.9	630	2	Q72155_THET2	Q72155	thermus the
589	40	88.9	631	2	Q6EUW2_XENLA	Q6euw2	xenopus lae
590	40	88.9	632	2	Q6YXW2_ORYSA	Q6yxw2	oryza sativ
591	40	88.9	633	2	Q7XI88_ORYSA	Q7xi88	oryza sativ
592	40	88.9	635	2	Q9WHJ7_FRMCV	Q9whj7	friend mink
593	40	88.9	636	1	ENV_MCFf3	P15073	mink cell f
594	40	88.9	636	2	Q6WP40_9GAMR	Q6wp40	murine leuk
595	40	88.9	636	2	Q6WP41_9GAMR	Q6wp41	murine leuk
596	40	88.9	636	2	Q6WP42_9GAMR	Q6wp42	murine leuk
597	40	88.9	636	2	Q6WP43_9GAMR	Q6wp43	murine leuk
598	40	88.9	636	2	Q85506_9GAMR	Q85506	murine leuk
599	40	88.9	638	2	Q9DQ21_9GAMR	Q9dq21	murine leuk
600	40	88.9	639	2	Q6T3B1_9GAMR	Q6t3b1	murine leuk
601	40	88.9	640	1	ENV_MCFf3	P03388	mink cell f
602	40	88.9	640	1	ENV_RMCFV	P06445	rauscher mi
603	40	88.9	640	2	Q9WHV5_9GAMR	Q9whv5	murine leuk
604	40	88.9	641	2	Q6WP47_MUSSP	Q6wp47	mus spretus
605	40	88.9	641	2	Q78N71_MOUSE	Q78n71	mus musculus
606	40	88.9	641	2	Q78N73_MOUSE	Q78n73	mus musculus
607	40	88.9	641	2	Q78N94_MOUSE	Q78n94	mus musculus
608	40	88.9	641	2	Q80SW7_MOUSE	Q80sw7	mus musculus
609	40	88.9	641	2	Q80SW8_MOUSE	Q80sw8	mus musculus
610	40	88.9	641	2	Q80SY0_MOUSE	Q80sy0	mus musculus
611	40	88.9	641	2	Q80SY1_MOUSE	Q80sy1	mus musculus
612	40	88.9	641	2	Q811M9_MOUSE	Q81lm9	mus musculus
613	40	88.9	641	2	Q811N4_MOUSE	Q81ln4	mus musculus
614	40	88.9	641	2	Q83363_9GAMR	Q83363	murine leuk
615	40	88.9	644	2	Q5RKN8_MOUSE	Q5rkn8	mus musculus

616	40	88.9	645	2	Q9H9W9_HUMAN	Q9h9w9	homo sapien
617	40	88.9	646	2	Q4HY89_GIBZE	Q4hy89	gibberella
618	40	88.9	653	2	Q73TB8_MYCPA	Q73tb8	mycobacteri
619	40	88.9	653	2	Q810W5_SPETR	Q810w5	spermophilu
620	40	88.9	655	1	FOX01_HUMAN	Q12778	homo sapien
621	40	88.9	655	2	Q5VYC7_HUMAN	Q5vyc7	homo sapien
622	40	88.9	655	2	Q8YMK1_ANASP	Q8ymk1	anabaena sp
623	40	88.9	658	2	Q8MQZ8_DROME	Q8mqz8	drosophila
624	40	88.9	669	1	MMP15_HUMAN	P51511	homo sapien
625	40	88.9	673	2	Q54E44_DICDI	Q54e44	dictyosteli
626	40	88.9	678	2	Q870Q4_NEUCR	Q870q4	neurospora
627	40	88.9	683	2	Q8LIW1_ORYSA	Q8liw1	oryza sativ
628	40	88.9	692	2	Q4QAH4_LEIMA	Q4qah4	leishmania
629	40	88.9	693	2	Q4T7R3_TETNG	Q4t7r3	tetraodon n
630	40	88.9	696	2	Q60YB5_CABBR	Q60yb5	caenorhabdi
631	40	88.9	701	2	Q7Z287_APLCA	Q7z287	aplysia cal
632	40	88.9	706	2	Q9VNZ5_DROME	Q9vnz5	drosophila
633	40	88.9	707	2	Q5NBU0_ORYSA	Q5nbu0	oryza sativ
634	40	88.9	708	2	Q9LJG1_ARATH	Q9ljg1	arabidopsis
635	40	88.9	710	2	Q9PYW6_GVXN	Q9pyw6	xestia c-ni
636	40	88.9	712	1	ZBT39_HUMAN	O15060	homo sapien
637	40	88.9	714	2	Q4I754_GIBZE	Q4i754	gibberella
638	40	88.9	717	2	Q5AMH8_CANAL	Q5amh8	candida alb
639	40	88.9	718	2	Q4TAL3_TETNG	Q4tal3	tetraodon n
640	40	88.9	720	2	Q4I4D0_GIBZE	Q4i4d0	gibberella
641	40	88.9	721	2	Q8N8Y6_HUMAN	Q8n8y6	homo sapien
642	40	88.9	721	2	Q8N3Y3_HUMAN	Q8n3y3	homo sapien
643	40	88.9	734	2	Q869R9_DICDI	Q869r9	dictyosteli
644	40	88.9	737	2	Q5NBD4_ORYSA	Q5nbd4	oryza sativ
645	40	88.9	740	2	Q59U10_CANAL	Q59u10	candida alb
646	40	88.9	742	2	Q59U34_CANAL	Q59u34	candida alb
647	40	88.9	744	2	Q4RMG0_TETNG	Q4rmg0	tetraodon n
648	40	88.9	751	1	PASTA_MOUSE	Q8biv7	mus musculu
649	40	88.9	751	1	PASTA_RAT	Q8k4s3	rattus norv
650	40	88.9	751	2	Q566E3_RAT	Q566e3	rattus norv
651	40	88.9	756	2	O18284_CAEEL	O18284	caenorhabdi
652	40	88.9	767	2	Q7ZVZ3_BRARE	Q7zvz3	brachydanio
653	40	88.9	768	2	Q5AFN4_CANAL	Q5afn4	candida alb
654	40	88.9	768	2	Q4WZK4_ASPFU	Q4wzk4	aspergillus
655	40	88.9	773	2	Q9ULJ9_DROME	Q9ulj9	drosophila
656	40	88.9	775	2	Q8RWW9_ARATH	Q8rww9	arabidopsis
657	40	88.9	781	2	Q4SB67_TETNG	Q4sb67	tetraodon n
658	40	88.9	782	1	PASTA_HUMAN	Q9y2w3	homo sapien
659	40	88.9	785	2	P74189_SYNY3	P74189	synechocyst
660	40	88.9	786	2	Q4HUC8_GIBZE	Q4huc8	gibberella
661	40	88.9	787	2	Q9VEJ1_DROME	Q9vej1	drosophila
662	40	88.9	791	2	Q5ZM84_CHICK	Q5zm84	gallus gall
663	40	88.9	792	2	Q653C0_ORYSA	Q653c0	oryza sativ
664	40	88.9	792	2	Q6EUD5_ORYSA	Q6eud5	oryza sativ
665	40	88.9	795	2	Q504V9_HUMAN	Q504v9	homo sapien
666	40	88.9	802	2	Q6GPA9_XENLA	Q6gpa9	xenopus lae
667	40	88.9	804	2	Q9UPS6_HUMAN	Q9ups6	homo sapien
668	40	88.9	804	2	Q4RX49_TETNG	Q4rx49	tetraodon n
669	40	88.9	805	2	Q61PW9_CABBR	Q61pw9	caenorhabdi
670	40	88.9	812	2	Q8LJ56_ORYSA	Q8lj56	oryza sativ
671	40	88.9	815	2	Q61EP1_CABBR	Q61ep1	caenorhabdi
672	40	88.9	826	1	LOZEN_DROME	Q9w349	drosophila
673	40	88.9	831	2	Q7XPI3_ORYSA	Q7xpi3	oryza sativ
674	40	88.9	834	2	Q5ACZ0_CANAL	Q5acz0	candida alb
675	40	88.9	835	2	Q6PGC9_MOUSE	Q6pgc9	mus musculu
676	40	88.9	839	2	Q9SN46_ARATH	Q9sn46	arabidopsis
677	40	88.9	840	2	Q57WS7_9TRYP	Q57ws7	trypanosoma
678	40	88.9	847	2	Q6AU77_ORYSA	Q6au77	oryza sativ
679	40	88.9	849	2	Q520N2_MAGGR	Q520n2	magnaporth
680	40	88.9	850	1	SEMLA_DROME	Q24322	drosophila
681	40	88.9	851	2	Q4P0G7_USTMA	Q4p0g7	ustilago ma
682	40	88.9	854	2	Q4WP55_ASPFU	Q4wp55	aspergillus
683	40	88.9	854	2	Q5JXM8_HUMAN	Q5jxm8	homo sapien
684	40	88.9	855	2	Q80TK9_MOUSE	Q80tk9	mus musculu
685	40	88.9	856	2	Q6ZQD1_MOUSE	Q6zqd1	mus musculu
686	40	88.9	859	2	Q4QG27_LEIMA	Q4gg27	leishmania
687	40	88.9	869	2	Q4S495_TETNG	Q4s495	tetraodon n
688	40	88.9	875	2	O48930_SOYBN	O48930	glycine max

689	40	88.9	890	2	Q9SYR0_SOYBN	Q9syro	glycine max	762	40	88.9	1239	2	Q9V4P1_DROME	Q9v4p1	drosophila
690	40	88.9	898	2	Q4IMV1_GIBZE	Q4imv1	gibberella	763	40	88.9	1240	2	P79773_CHICK	P79773	gallus gall
691	40	88.9	899	2	Q8N304_HUMAN	Q8n304	homo sapien	764	40	88.9	1242	1	IRS1_HUMAN	IRS1	homo sapien
692	40	88.9	905	2	Q7XPU8_ORYSA	Q7xpu8	oryza sativ	765	40	88.9	1267	2	Q98945_CHICK	Q98945	gallus gall
693	40	88.9	911	2	Q8CGD1_MOUSE	Q8cgdl	mus musculus	766	40	88.9	1283	2	Q8T5H0_ANOGA	Q8t5h0	anopheles g
694	40	88.9	917	2	Q8CFQ8_MOUSE	Q8cfq8	mus musculus	767	40	88.9	1316	2	Q75128_HUMAN	Q75128	homo sapien
695	40	88.9	919	2	Q6NZL0_MOUSE	Q6nzl0	mus musculus	768	40	88.9	1351	1	OVO_DROME	P51521	drosophila
696	40	88.9	932	2	Q54SG9_DICDI	Q54sg9	dictyosteli	769	40	88.9	1374	2	Q6AWI5_DROME	Q6awi5	drosophila
697	40	88.9	938	2	Q96VI4_PNECA	Q96vi4	pneumocysti	770	40	88.9	1388	2	Q9BYU8_HUMAN	Q9byu8	homo sapien
698	40	88.9	947	2	Q5TF21_HUMAN	Q5tf21	homo sapien	771	40	88.9	1396	2	Q4QJB8_LEIMA	Q4qjb8	leishmania
699	40	88.9	950	2	Q4P2X1_USTMA	Q4p2x1	ustilago ma	772	40	88.9	1403	1	PROS_DROME	P29617	drosophila
700	40	88.9	954	2	Q60TB3_CABBR	Q60tb3	caenorhabdi	773	40	88.9	1425	2	Q4P427_USTMA	Q4p427	ustilago ma
701	40	88.9	961	2	Q51T63_MAGGR	Q51t63	magnaporthe	774	40	88.9	1426	2	O15837_LEIMA	O15837	leishmania
702	40	88.9	963	1	RFX1_MOUSE	P48377	mus musculus	775	40	88.9	1437	2	Q9BYU9_HUMAN	Q9byu9	homo sapien
703	40	88.9	963	2	Q54F88_DICDI	Q54f88	dictyosteli	776	40	88.9	1437	2	Q9BZ95_HUMAN	Q9bz95	homo sapien
704	40	88.9	963	2	Q8BJK3_MOUSE	Q8bjk3	mus musculus	777	40	88.9	1442	2	Q7QIY3_ANOGA	Q7qiY3	anopheles g
705	40	88.9	963	2	Q6PGH8_MOUSE	Q6pgh8	mus musculus	778	40	88.9	1458	2	Q5IU00_MAGGR	Q5lu00	magnaporthe
706	40	88.9	968	2	Q9HCM7_HUMAN	Q9hcm7	homo sapien	779	40	88.9	1500	2	Q7XWU1_ORYSA	Q7xwu1	oryza sativ
707	40	88.9	972	2	Q51Z25_MAGGR	Q51z25	magnaporthe	780	40	88.9	1581	2	Q6PDJ2_MOUSE	Q6pdj2	mus musculus
708	40	88.9	980	2	Q5ASA5_EMENI	Q5asa5	aspergillus	781	40	88.9	1601	2	Q5K9V7_CRYNE	Q5k9v7	cryptococcu
709	40	88.9	987	2	Q59VG5_CANAL	Q59vg5	candida alb	782	40	88.9	1604	2	Q55JP5_CRYNE	Q55jp5	cryptococcu
710	40	88.9	987	2	Q59VK1_CANAL	Q59vk1	candida alb	783	40	88.9	1609	2	Q7RYS2_NEUCR	Q7rys2	neurospora
711	40	88.9	994	2	Q9V4S5_DROME	Q9v4s5	drosophila	784	40	88.9	1715	2	Q4Q7H2_LEIMA	Q4q7h2	leishmania
712	40	88.9	995	2	Q4IQ76_GIBZE	Q4iq76	gibberella	785	40	88.9	1768	2	Q7S8D6_NEUCR	Q7s8d6	neurospora
713	40	88.9	996	2	Q5CKJ5_CRYHO	Q5ckj5	cryptospori	786	40	88.9	1840	2	Q5SRY3_MOUSE	Q5sry3	mus musculus
714	40	88.9	999	2	Q4SS78_TETNG	Q4ss78	tetraodon n	787	40	88.9	1846	2	Q86YS8_HUMAN	Q86ys8	homo sapien
715	40	88.9	1003	2	Q4WXI4_ASPFU	Q4wx14	aspergillus	788	40	88.9	1848	2	Q69ZN8_MOUSE	Q69zn8	mus musculus
716	40	88.9	1004	2	Q4QHS6_LEIMA	Q4qhs6	leishmania	789	40	88.9	1889	1	RAI1_MOUSE	Q61818	mus musculus
717	40	88.9	1023	2	Q53Q63_ORYSA	Q53q63	oryza sativ	790	40	88.9	1889	2	Q5QGZ6_MOUSE	Q5qgz6	mus musculus
718	40	88.9	1024	2	Q4WXM8_ASPFU	Q4wxm8	aspergillus	791	40	88.9	1889	2	Q5SRY2_MOUSE	Q5sry2	mus musculus
719	40	88.9	1026	1	STAU_DROME	P25159	drosophila	792	40	88.9	1906	1	RAI1_HUMAN	Q7z5j4	homo sapien
720	40	88.9	1036	2	Q7SD91_NEUCR	Q7sd91	neurospora	793	40	88.9	1907	2	Q5ICW5_BRARE	Q5icw5	brachydanio
721	40	88.9	1048	2	Q5EBL2_HUMAN	Q5eb12	homo sapien	794	40	88.9	1949	2	Q8UZB6_9VIRU	Q8uzb6	grapevine f
722	40	88.9	1056	2	Q6PDK1_MOUSE	Q6pdk1	mus musculus	795	40	88.9	1966	2	Q9NHX6_DROME	Q9nhx6	drosophila
723	40	88.9	1057	2	Q9EQN4_MOUSE	Q9eqn4	mus musculus	796	40	88.9	1966	2	Q8IQA6_DROME	Q8iqa6	drosophila
724	40	88.9	1058	2	Q69Z44_MOUSE	Q69z44	mus musculus	797	40	88.9	1985	2	Q9VSK5_DROME	Q9vsk5	drosophila
725	40	88.9	1059	2	Q7S0G4_NEUCR	Q7s0g4	neurospora	798	40	88.9	1985	2	Q8T9N4_DROME	Q8t9n4	drosophila
726	40	88.9	1059	2	Q8C5X1_MOUSE	Q8c5x1	mus musculus	799	40	88.9	1985	2	Q7KUA8_DROME	Q7kua8	drosophila
727	40	88.9	1059	2	Q6PAL5_MOUSE	Q6pal5	mus musculus	800	40	88.9	1988	2	Q86BH2_DROME	Q86bh2	drosophila
728	40	88.9	1062	2	Q63ZY4_HUMAN	Q63zy4	homo sapien	801	40	88.9	1990	2	Q9U8Q0_DROME	Q9u8q0	drosophila
729	40	88.9	1063	2	Q6ZS14_HUMAN	Q6z814	homo sapien	802	40	88.9	1995	2	Q9W244_DROME	Q9w244	drosophila
730	40	88.9	1070	1	Y0355_HUMAN	O15063	homo sapien	803	40	88.9	2021	2	Q4RF83_TETNG	Q4rf83	tetraodon n
731	40	88.9	1070	2	Q5R4J3_PONPY	Q5r4j3	pongo pygma	804	40	88.9	2053	2	Q4QDU5_LEIMA	Q4qdu5	leishmania
732	40	88.9	1070	2	Q5REE5_PONPY	Q5ree5	pongo pygma	805	40	88.9	2075	2	Q9VXY2_DROME	Q9vxy2	drosophila
733	40	88.9	1070	2	Q4R3A1_MACFA	Q4r3a1	macaca faec	806	40	88.9	2185	2	Q75MN6_HUMAN	Q75mn6	homo sapien
734	40	88.9	1075	1	ATX2L_HUMAN	Q8wmm7	homo sapien	807	40	88.9	2220	2	Q4PBQ9_USTMA	Q4pbq9	ustilago ma
735	40	88.9	1078	2	Q59MQ9_CANAL	Q59mq9	candida alb	808	40	88.9	2246	2	Q6J514_BRARE	Q6j514	brachydanio
736	40	88.9	1078	2	Q8J2J5_CANAL	Q8j2j5	candida alb	809	40	88.9	2274	2	Q54W13_DICDI	Q54w13	dictyosteli
737	40	88.9	1085	2	Q4QHD4_LEIMA	Q4qhd4	leishmania	810	40	88.9	2343	2	Q4QIK3_LEIMA	Q4qik3	leishmania
738	40	88.9	1092	2	Q8CFT2_MOUSE	Q8cft2	mus musculus	811	40	88.9	2403	2	Q68CP0_HUMAN	Q68cp0	homo sapien
739	40	88.9	1094	2	Q5B395_EMENI	Q5b395	aspergillus	812	40	88.9	2476	2	Q4RPG5_TETNG	Q4rpg5	tetraodon n
740	40	88.9	1096	2	Q86PF1_BACOL	Q86pf1	bactrocera	813	40	88.9	2506	2	Q63HJ5_HUMAN	Q63hj5	homo sapien
741	40	88.9	1105	1	DPOD1_ORYSA	Q9lre6	oryza sativ	814	40	88.9	2563	2	Q54LJ4_DICDI	Q54lj4	dictyosteli
742	40	88.9	1105	2	Q53P50_ORYSA	Q53p50	oryza sativ	815	40	88.9	2592	2	Q9PJ30_NEUCR	Q9pj30	neurospora
743	40	88.9	1117	2	Q8NCN1_HUMAN	Q8ncn1	homo sapien	816	40	88.9	2643	2	Q4IIG9_GIBZE	Q4ilg9	gibberella
744	40	88.9	1132	2	Q86Y26_HUMAN	Q86y26	homo sapien	817	40	88.9	2646	2	Q8I220_PLAF7	Q8i220	plasmodium
745	40	88.9	1134	2	Q96JH1_HUMAN	Q96jh1	homo sapien	818	40	88.9	2938	2	Q9VEE5_DROME	Q9vee5	drosophila
746	40	88.9	1146	2	Q86XA7_HUMAN	Q86xa7	homo sapien	819	40	88.9	3072	1	EP400_MOUSE	Q8chi8	mus musculus
747	40	88.9	1153	2	Q8MQW5_DROME	Q8mqw5	drosophila	820	40	88.9	3110	1	HD_RAT	P51111	rattus norv
748	40	88.9	1158	2	Q7PTX0_ANOGA	Q7ptx0	anopheles g	821	40	88.9	3119	2	Q8IHM0_PLAF7	Q8ihm0	plasmodium
749	40	88.9	1160	2	Q4WXM4_ASPFU	Q4wxm4	aspergillus	822	40	88.9	3139	2	Q9GM99_PIG	Q9gm99	sus scrofa
750	40	88.9	1162	2	Q867K1_BACOL	Q867k1	bactrocera	823	40	88.9	3146	2	Q9VUB5_DROME	Q9vub5	drosophila
751	40	88.9	1164	2	Q6CH17_YARLI	Q6ch17	yarrowia li	824	40	88.9	3246	2	Q4QGR1_LEIMA	Q4qgr1	leishmania
752	40	88.9	1170	2	Q95TI5_DROME	Q95ti5	drosophila	825	40	88.9	3396	2	Q5YLV9_MOUSE	Q5ylv9	mus musculus
753	40	88.9	1174	1	KPC1_COCHE	Q42632	cochliobolu	826	40	88.9	3889	2	Q6SSE8_CHLRE	Q6sse8	chlamydomon
754	40	88.9	1181	2	Q7SEP4_NEUCR	Q7sep4	neurospora	827	39	86.7	39	2	Q9BR93_HUMAN	Q9br93	homo sapien
755	40	88.9	1185	2	Q53PZ5_ORYSA	Q53pz5	oryza sativ	828	39	86.7	45	2	Q8W701_9CAUD	Q8w701	cyanophage
756	40	88.9	1201	2	Q24240_DROME	Q24240	drosophila	829	39	86.7	62	2	Q6Z053_ORYSA	Q6z053	oryza sativ
757	40	88.9	1211	2	Q69NX2_ORYSA	Q69nx2	oryza sativ	830	39	86.7	72	2	Q4K9D4_PSEF5	Q4k9d4	pseudomonas
758	40	88.9	1212	2	Q8MKN2_DROME	Q8mkn2	drosophila	831	39	86.7	77	2	Q9NRK3_HUMAN	Q9nrk3	homo sapien
759	40	88.9	1214	1	DGKD_HUMAN	Q16760	homo sapien	832	39	86.7	82	2	Q4TPY1_9SPHN	Q4tpy1	erythro bact
760	40	88.9	1223	2	Q54W89_DICDI	Q54w89	dictyosteli	833	39	86.7	90	2	Q4VMJ0_MANSE	Q4vmj0	manduca sex
761	40	88.9	1239	2	O01505_CABEL	O01505	caenorhabdi	834	39	86.7	94	2	Q8W6Y8_9CAUD	Q8w6y8	cyanophage

835	39	86.7	95	2	Q8LBN7_ARATH	Q8lbn7 arabidopsis
836	39	86.7	104	2	Q9FW1_ORYSA	Q9fww1 oryza sativ
837	39	86.7	106	2	Q8L522_ORYSA	Q8l522 oryza sativ
838	39	86.7	108	2	Q91AX2_9PAPI	Q9lax2 human papil
839	39	86.7	111	2	Q6YTM7_ORYSA	Q6ytm7 oryza sativ
840	39	86.7	116	2	Q5VMR4_ORYSA	Q5vmr4 oryza sativ
841	39	86.7	117	2	Q57Y88_9TRYP	Q57y88 trypanosoma
842	39	86.7	117	2	Q7XHN1_ORYSA	Q7xhn1 oryza sativ
843	39	86.7	118	2	Q5NB16_ORYSA	Q5nb16 oryza sativ
844	39	86.7	123	2	Q4NW92_9DELT	Q4nw92 anaeromyxob
845	39	86.7	126	2	Q9LPK9_ARATH	Q9lpk9 arabidopsis
846	39	86.7	128	2	Q9AV28_ORYSA	Q9av28 oryza sativ
847	39	86.7	129	2	Q6K9A5_ORYSA	Q6k9a5 oryza sativ
848	39	86.7	130	2	Q842K3_ECOLI	Q842k3 escherichia
849	39	86.7	131	2	Q99PH8_MOUSE	Q99ph8 mus musculus
850	39	86.7	133	2	Q7SEW5_NEUCR	Q7sew5 neurospora
851	39	86.7	134	2	Q4RFZ9_TETNG	Q4rfz9 tetraodon n
852	39	86.7	136	2	Q6Z268_ORYSA	Q6z268 oryza sativ
853	39	86.7	136	2	Q9WNX1_9DELA	Q9wnx1 human t-lym
854	39	86.7	137	2	Q05284_ECOLI	Q05284 escherichia
855	39	86.7	137	2	Q5SPH8_BRARE	Q5sph8 brachydanio
856	39	86.7	140	2	Q6MUW0_NEUCR	Q6muw0 neurospora
857	39	86.7	147	1	SMR1_MOUSE	Q61900 mus musculus
858	39	86.7	148	2	Q7EYB6_ORYSA	Q7eyb6 oryza sativ
859	39	86.7	148	2	Q9RS96_DEIRA	Q9rs96 deinococcus
860	39	86.7	148	2	Q78DZ0_RATRT	Q78dz0 rattus ratt
861	39	86.7	148	2	Q64371_RAT	Q64371 rattus norv
862	39	86.7	153	2	Q653R3_ORYSA	Q653r3 oryza sativ
863	39	86.7	159	2	Q5W6Y2_ORYSA	Q5w6y2 oryza sativ
864	39	86.7	160	2	Q9AJ06_ECOLI	Q9aj06 escherichia
865	39	86.7	161	2	Q9M0N0_ARATH	Q9m0n0 arabidopsis
866	39	86.7	161	2	Q81881_ARATH	Q81881 arabidopsis
867	39	86.7	163	2	Q7V6A1_PROMM	Q7v6a1 prochloroco
868	39	86.7	164	2	Q9AUR6_ORYSA	Q9aur6 oryza sativ
869	39	86.7	166	2	Q6Z4S5_ORYSA	Q6z4s5 oryza sativ
870	39	86.7	170	2	Q52584_ECOLI	Q52584 escherichia
871	39	86.7	172	2	Q9DCH0_MOUSE	Q9dch0 mus musculus
872	39	86.7	174	2	Q7VPW5_CHLPN	Q7vpw5 chlamydia p
873	39	86.7	178	2	Q75AN4_ASHGO	Q75an4 ashbya goss
874	39	86.7	179	2	Q9D0Y6_MOUSE	Q9d0y6 mus musculus
875	39	86.7	180	2	Q5JNH2_ORYSA	Q5jnh2 oryza sativ
876	39	86.7	181	2	Q35328_MOUSE	Q35328 mus musculus
877	39	86.7	182	2	Q4YEI4_PLABE	Q4yei4 plasmodium
878	39	86.7	182	2	Q8CEB2_MOUSE	Q8ceb2 mus musculus
879	39	86.7	185	2	Q53R94_HUMAN	Q53r94 homo sapien
880	39	86.7	188	2	Q51Q52_MAGGR	Q51q52 magnaporthe
881	39	86.7	189	2	Q67V82_ORYSA	Q67v82 oryza sativ
882	39	86.7	190	2	Q6IL94_DROME	Q6il94 drosophila
883	39	86.7	192	2	Q51SS5_MAGGR	Q51ss5 magnaporthe
884	39	86.7	195	2	Q8S5W0_ORYSA	Q8s5w0 oryza sativ
885	39	86.7	196	2	Q8C0B8_MOUSE	Q8c0b8 mus musculus
886	39	86.7	198	2	Q7XCS7_ORYSA	Q7xc87 oryza sativ
887	39	86.7	198	2	Q94HY7_ORYSA	Q94hy7 oryza sativ
888	39	86.7	198	2	Q9S2J6_STRCO	Q9s2j6 streptomyce
889	39	86.7	198	2	Q5TYZ7_BRARE	Q5tyz7 brachydanio
890	39	86.7	199	2	Q57YJ3_9TRYP	Q57yj3 trypanosoma
891	39	86.7	199	2	Q6YU08_ORYSA	Q6yuu8 oryza sativ
892	39	86.7	202	2	Q82444_9DELA	Q82444 human t-lym
893	39	86.7	204	2	Q5N7R3_ORYSA	Q5n7r3 oryza sativ
894	39	86.7	204	2	Q5WME5_ECOLI	Q5wme5 escherichia
895	39	86.7	204	2	Q9CUA2_MOUSE	Q9cua2 mus musculus
896	39	86.7	206	2	Q8GRR7_ORYSA	Q8grr7 oryza sativ
897	39	86.7	206	2	Q52152_ECOLI	Q52152 escherichia
898	39	86.7	207	2	Q9ROE1_ECOLI	Q9rqe1 escherichia
899	39	86.7	208	2	Q9CUQ2_MOUSE	Q9cuq2 mus musculus
900	39	86.7	211	2	Q84PY7_ORYSA	Q84py7 oryza sativ
901	39	86.7	213	2	Q6Z0B3_ORYSA	Q6z0b3 oryza sativ
902	39	86.7	213	2	Q9Z7M6_CHLPN	Q9z7m6 chlamydia p
903	39	86.7	217	2	Q8S7F2_ORYSA	Q8s7f2 oryza sativ
904	39	86.7	219	2	Q9XWF3_CAEEL	Q9xwf3 caenorhabdi
905	39	86.7	219	2	Q9M3G8_ARATH	Q9m3g8 arabidopsis
906	39	86.7	219	2	Q994D5_9ADEN	Q994d5 porcine ade
907	39	86.7	220	2	Q96030_HUMAN	Q96030 homo sapien

908	39	86.7	220	2	Q5T4W7_HUMAN	Q5t4w7 homo sapien
909	39	86.7	221	2	Q6Z070_ORYSA	Q6z070 oryza sativ
910	39	86.7	224	2	Q6L4Q0_ORYSA	Q6l4q0 oryza sativ
911	39	86.7	225	2	Q6N0Q5_RHOPA	Q6n0q5 rhodopseudo
912	39	86.7	228	2	Q6P6A3_HUMAN	Q6p6a3 homo sapien
913	39	86.7	231	2	Q4ILQ0_GIBZE	Q4ilq0 gibberella
914	39	86.7	231	2	Q8N9Z2_HUMAN	Q8n9z2 homo sapien
915	39	86.7	232	2	Q6H694_ORYSA	Q6h694 oryza sativ
916	39	86.7	233	2	Q679P4_MOUSE	Q679p4 mus musculus
917	39	86.7	234	2	Q8H0A9_ORYSA	Q8h0a9 oryza sativ
918	39	86.7	235	2	Q6ZBX1_ORYSA	Q6zbx1 oryza sativ
919	39	86.7	236	1	UL51_PRVKA	Q85227 pseudorabie
920	39	86.7	236	2	Q5PP96_9ALPH	Q5pp96 suid herpes
921	39	86.7	237	2	Q95441_HUMAN	Q95441 homo sapien
922	39	86.7	241	2	Q7NDI6_GLOVI	Q7ndi6 gloeobacter
923	39	86.7	242	2	Q5Z4Q0_ORYSA	Q5z4q0 oryza sativ
924	39	86.7	244	2	Q6ETA6_ORYSA	Q6eta6 oryza sativ
925	39	86.7	244	2	Q4NQD2_9DELT	Q4nqd2 anaeromyxob
926	39	86.7	246	2	Q6ZYK2_9VIRU	Q6zyk2 pyrobaculum
927	39	86.7	247	2	Q5N9K0_ORYSA	Q5n9k0 oryza sativ
928	39	86.7	248	2	Q85622_ECOLI	Q85622 escherichia
929	39	86.7	248	2	Q7DB85_ECO57	Q7db85 escherichia
930	39	86.7	256	2	Q5YB71_MANSE	Q5yb71 manduca sex
931	39	86.7	259	2	Q7S5M0_NEUCR	Q7s5m0 neurospora
932	39	86.7	259	2	Q5SND1_ORYSA	Q5snd1 oryza sativ
933	39	86.7	260	2	Q850Z3_ORYSA	Q850z3 oryza sativ
934	39	86.7	261	2	Q5AK47_CANAL	Q5ak47 candida alb
935	39	86.7	264	2	Q9UPA8_HUMAN	Q9upa8 homo sapien
936	39	86.7	267	2	Q7XXM9_ORYSA	Q7xxm9 oryza sativ
937	39	86.7	270	2	Q4NQRI_9DELT	Q4nqri anaeromyxob
938	39	86.7	271	2	Q4S631_TETNG	Q4s631 tetraodon n
939	39	86.7	275	2	Q654Z7_ORYSA	Q654z7 oryza sativ
940	39	86.7	278	1	TNFL6_RAT	Q36940 rattus norv
941	39	86.7	278	2	Q9BH93_DROME	Q9bh93 drosophila
942	39	86.7	279	1	TNFL6_MOUSE	Q41047 mus musculus
943	39	86.7	279	2	Q7TMV9_MOUSE	Q7tmv9 mus musculus
944	39	86.7	279	2	Q544E9_MOUSE	Q544e9 mus musculus
945	39	86.7	280	1	TNFL6_CERTO	Q9bdsn1 c tumor nec
946	39	86.7	280	1	TNFL6_MACFA	Q63308 m tumor nec
947	39	86.7	280	1	TNFL6_MACMU	Q63307 m tumor nec
948	39	86.7	280	1	TNFL6_MACNE	Q63306 m tumor nec
949	39	86.7	280	2	Q861W5_FELCA	Q861w5 felis silve
950	39	86.7	281	1	TNFL6_HUMAN	Q48023 h tumor nec
951	39	86.7	281	2	Q53ZZ1_HUMAN	Q53zz1 homo sapien
952	39	86.7	282	1	TNFL6_PIG	Q9bea8 sus scrofa
953	39	86.7	286	2	Q9VH99_DROME	Q9vh99 drosophila
954	39	86.7	288	2	Q61TJ4_CAEBR	Q6ltj4 caenorhabdi
955	39	86.7	291	2	Q51LQ6_MAGGR	Q51lq6 magnaporthe
956	39	86.7	291	2	Q53LV0_ORYSA	Q53lv0 oryza sativ
957	39	86.7	292	2	Q96881_DROME	Q96881 drosophila
958	39	86.7	294	2	Q5K5R0_ECOLI	Q5k5r0 escherichia
959	39	86.7	294	2	Q4TQZ7_9SPHN	Q4tqz7 erythrobact
960	39	86.7	295	2	Q94DE7_ORYSA	Q94de7 oryza sativ
961	39	86.7	296	2	Q95R37_DROME	Q95r37 drosophila
962	39	86.7	296	2	Q9VMI6_DROME	Q9vmi6 drosophila
963	39	86.7	302	2	Q86UI6_HUMAN	Q86ul6 homo sapien
964	39	86.7	306	2	P93845_PEA	Q93845 pisum sativ
965	39	86.7	306	2	Q6ENK1_ORYSA	Q6enk1 oryza sativ
966	39	86.7	308	2	Q4S667_TETNG	Q4s667 tetraodon n
967	39	86.7	310	2	Q75K33_ORYSA	Q75k33 oryza sativ
968	39	86.7	313	2	Q7VX28_BORPE	Q7vx28 bordetella
969	39	86.7	313	2	Q7W7Y4_BORPA	Q7w7y4 bordetella
970	39	86.7	313	2	Q7WLC4_BORBR	Q7wlc4 bordetella
971	39	86.7	314	1	PITX1_HUMAN	Q78337 homo sapien
972	39	86.7	314	2	Q41974_MHV68	Q41974 murid herpe
973	39	86.7	315	1	PITX1_MOUSE	Q70314 mus musculus
974	39	86.7	315	1	PITX1_RAT	Q99na7 rattus norv
975	39	86.7	316	2	Q6ZHH2_ORYSA	Q6zhh2 oryza sativ
976	39	86.7	318	2	Q4IL55_GIBZE	Q4il55 gibberella
977	39	86.7	322	2	Q851R4_ORYSA	Q851r4 oryza sativ
978	39	86.7	326	2	Q599K1_BOVIN	Q599k1 bos taurus
979	39	86.7	327	2	Q99JK6_MOUSE	Q99jk6 mus musculus
980	39	86.7	329	2	Q96S04_HUMAN	Q96s04 homo sapien

981 39 86.7 329 2 Q5ZDM9_ORYSA Q5zdm9 oryza sativ
982 39 86.7 333 1 HME1_CHICK Q05916 gallus gall
983 39 86.7 333 2 Q6Z4H6_ORYSA Q6z4h6 oryza sativ
984 39 86.7 340 2 Q6Z303_ORYSA Q6z303 oryza sativ
985 39 86.7 343 2 Q6IPN0_HUMAN Q6ipn0 homo sapien
986 39 86.7 345 2 Q6NUL8_HUMAN Q6nul8 homo sapien
987 39 86.7 346 2 Q75525_HUMAN Q75525 homo sapien
988 39 86.7 346 2 Q88624_MOUSE Q88624 mus musculu
989 39 86.7 346 2 Q9JLPI_RAT Q9jlp1 rattus norv
990 39 86.7 346 2 Q9R226_MOUSE Q9r226 mus musculu
991 39 86.7 347 2 Q6K911_ORYSA Q6k911 oryza sativ
992 39 86.7 348 2 Q7XTQ3_ORYSA Q7xtq3 oryza sativ
993 39 86.7 348 2 Q9M1F3_ARATH Q9m1f3 arabidopsis
994 39 86.7 350 2 Q94BW2_ARATH Q94bw2 arabidopsis
995 39 86.7 351 2 Q94GL5_ORYSA Q94gl5 oryza sativ
996 39 86.7 351 2 Q7BQA4_MYCPH Q7bqa4 mycobacteri
997 39 86.7 354 2 Q84QZ0_ORYSA Q84qz0 oryza sativ
998 39 86.7 354 2 Q94LH4_ORYSA Q94lh4 oryza sativ
999 39 86.7 355 2 Q9LGY9_ORYSA Q9lgy9 oryza sativ
1000 39 86.7 359 2 Q50Q27_ENTHI Q50q27 entamoeba h

ALIGNMENTS

RESULT 1
Q6IIP9_DROME
ID Q6IIP9_DROME PRELIMINARY; PRT; 66 AA.
AC Q6IIP9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE HDC17385.

GN ORFNames=HDC17385;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;

RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14709175; DOI=10.1186/gb-2003-5-1-r3;
RA Hild M., Beckmann B., Haas S.A., Koch B., Solovyev V., Busold C.,
RA Fellenberg K., Boutros M., Vingron M., Sauer F., Hoheisel J.D.,
RA Paro R.;
RT "An integrated gene annotation and transcriptional profiling approach
RT towards the full gene content of the Drosophila genome."
RL Genome Biol. 5:RESEARCH0003.1-RESEARCH0003.17(2003).
CC -|- MISCELLANEOUS: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ third party annotation (TPA) entry.
DR EMBL; BK003017; DAA03217.1; -; Genomic DNA.
SQ SEQUENCE 66 AA; 7652 MW; D0C24FDE4E24C1D3 CRC64;

Query Match 100.0%; Score 45; DB 2; Length 66;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 49 RPQPPPP 55

RESULT 2
Q5VP16_ORYSA
ID Q5VP16_ORYSA PRELIMINARY; PRT; 73 AA.
AC Q5VP16;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein OSJNBa0091E23.6 (Hypothetical protein
DE B1096D03.50).
GN Name=OSJNBa0091E23.6; Synonyms=B1096D03.50;
OS Oryza sativa (japonica cultivar-group).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Niumura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
RA Ikeno M., Itoh S., Itoh T., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,
RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA Zhong H., Iwana H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA Yano M., Jiang J., Gojobori T.;
RT "The genome sequence and structure of rice chromosome 1.";
RL Nature 420:312-316(2002).
DR EMBL; AP003853; BAD68808.1; -; Genomic DNA.
DR EMBL; AP003536; BAD68493.1; -; Genomic DNA.
KW Hypothetical protein.
SQ SEQUENCE 73 AA; 7805 MW; 5E2E35B132B8BAB4 CRC64;

Query Match 100.0%; Score 45; DB 2; Length 73;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 33 RPQPPPP 39

RESULT 3
Q5VPA7_ORYSA
ID Q5VPA7_ORYSA PRELIMINARY; PRT; 119 AA.
AC Q5VPA7;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein P0498C03.3 (Hypothetical protein
DE P0566A10.37).
GN Name=P0498C03.3; Synonyms=P0566A10.37;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC
RT clone:P0498C03.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC
RT clone:P0566A10.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003724; BAD68718.1; -; Genomic DNA.
DR EMBL; AP003630; BAD68611.1; -; Genomic DNA.
KW Hypothetical protein.
SQ SEQUENCE 119 AA; 12649 MW; D69766963F48B194 CRC64;

Query Match 100.0%; Score 45; DB 2; Length 119;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7

Db 96 RPQPPPP 102
|||||
RESULT 4
Q6LDK2_9MURI
ID Q6LDK2_9MURI PRELIMINARY; PRT; 125 AA.
AC Q6LDK2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Synapsin I (Fragment).
GN Name=SYN1;
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=90368667; PubMed=2118519;
RA Sauerwald A., Hoesche C., Oschwald R., Kilimann M.W.;
RT "The 5'-flanking region of the synapsin I gene. A G+C-rich, TATA- and
RT CAAT-less, phylogenetically conserved sequence with cell type-specific
RT promoter function."
RL J. Biol. Chem. 265:14932-14937(1990).
DR EMBL; M55300; AAA42199.1; -; Genomic DNA.
DR GO; GO:0008021; C:synaptic vesicle; IEA.
DR GO; GO:0007269; P:neurotransmitter secretion; IEA.
DR InterPro; IPR001359; Synapsin.
DR Pfam; PF02078; Synapsin N; 1.
DR PROSITE; PS00415; SYNAPSIN_1; 1.
FT NON TER 1
SQ SEQUENCE 125 AA; 12193 MW; DE1E6A61659EFCA9 CRC64;
Query Match 100.0%; Score 45; DB 2; Length 125;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RPQPPPP 7
|||||
Db 27 RPQPPPP 33
|||||
RESULT 5
Q6YZB3_ORYSA
ID Q6YZB3_ORYSA PRELIMINARY; PRT; 132 AA.
AC Q6YZB3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein B1099H05.16 (Hypothetical protein
DE P0020B10.40).
GN Name=B1099H05.16; Synonyms=P0020B10.40;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 8, BAC
RT clone:B1099H05."
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005531; BAD05764.1; -; Genomic DNA.
DR EMBL; AP004656; BAD03335.1; -; Genomic DNA.
DR Gramene; Q6YZB3; -.

KW Hypothetical protein.
SQ SEQUENCE 132 AA; 14128 MW; B943F2A5417B9CD3 CRC64;
Query Match 100.0%; Score 45; DB 2; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RPQPPPP 7
|||||
Db 47 RPQPPPP 53
|||||
RESULT 6
Q7F8E9_ORYSA
ID Q7F8E9_ORYSA PRELIMINARY; PRT; 165 AA.
AC Q7F8E9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein P0644B06.50 (Hypothetical protein
DE P0514G12.6).
GN Name=P0644B06.50; Synonyms=P0514G12.6;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC
RT clone:P0644B06."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC
RT clone:P0514G12."
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP001129; BAA90640.1; -; Genomic DNA.
DR EMBL; AP000616; BAD67680.1; -; Genomic DNA.
DR Gramene; Q7F8E9; -.
KW Hypothetical protein.
SQ SEQUENCE 165 AA; 16785 MW; 2E7CD76B9017B159 CRC64;
Query Match 100.0%; Score 45; DB 2; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RPQPPPP 7
|||||
Db 106 RPQPPPP 112
|||||
RESULT 7
Q654U7_ORYSA
ID Q654U7_ORYSA PRELIMINARY; PRT; 208 AA.
AC Q654U7;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein OJ1136_C11.22.
GN Name=OJ1136_C11.22;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, BAC
RT clone:OJ1136_C11."
RN [2]


```
RL Submitted (AUG-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AP004027; BAD45670.1; -; Genomic_DNA.
DR Gramene; Q654U7; -.
KW Hypothetical protein.
SQ SEQUENCE 208 AA; 21473 MW; 2C2EF7BC0E27DEED CRC64;

Query Match 100.0%; Score 45; DB 2; Length 208;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 88 RPQPPPP 94

RESULT 8
Q9ZA63 NEIME
ID Q9ZA63 NEIME PRELIMINARY; PRT; 270 AA.
AC Q9ZA63;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B1940;
RX MEDLINE=99005251; PubMed=9790590; DOI=10.1007/s004380050823;
RA Claus H.; Frosch M.; Vogel U.;
RT "Identification of a hotspot for transformation of Neisseria
RT meningitidis by shuttle mutagenesis using signature-tagged
RT transposons.";
RL Mol. Gen. Genet. 259:363-371(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B1940;
RA Claus H.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AJ000250; CAA03966.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 270 AA; 30192 MW; 612DE74D02BA09AA CRC64;

Query Match 100.0%; Score 45; DB 2; Length 270;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 116 RPQPPPP 122

RESULT 9
Q6Z7W0 ORYSA
ID Q6Z7W0 ORYSA PRELIMINARY; PRT; 291 AA.
AC Q6Z7W0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Peptide transporter-like.
GN Name=P0471A11.14;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T.; Matsumoto T.; Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 2, PAC
RT clone:P0471A11.";
```

```
RL Submitted (MAR-2002) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AP004814; BAD17084.1; -; Genomic_DNA.
DR Gramene; Q6Z7W0; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006857; P:oligopeptide transport; IEA.
DR InterPro; IPR000109; PTR2.
DR Pfam; PF00854; PTR2; 1.
SQ SEQUENCE 291 AA; 31319 MW; E08E8B7C4D83E591 CRC64;

Query Match 100.0%; Score 45; DB 2; Length 291;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 59 RPQPPPP 65

RESULT 10
Q8LIW2 ORYSA
ID Q8LIW2 ORYSA PRELIMINARY; PRT; 300 AA.
AC Q8LIW2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE P0497A05.11 protein.
GN Name=P0497A05.11;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22337376; PubMed=12447438; DOI=10.1038/nature01184;
RA Sasaki T.; Matsumoto T.; Yamamoto K.; Sakata K.; Baba T.; Katayose Y.;
RA Wu J.; Nimura Y.; Cheng Z.; Nagamura Y.; Antonio B.A.; Kanamori H.;
RA Hosokawa S.; Masukawa M.; Arikawa K.; Chiden Y.; Hayashi M.;
RA Okamoto M.; Ando T.; Aoki H.; Arita K.; Hamada M.; Harada C.;
RA Hijishita S.; Honda M.; Ichikawa Y.; Idonuma A.; Iijima M.; Ikeda M.;
RA Ikeno M.; Ito S.; Ito T.; Ito Y.; Ito Y.; Iwabuchi A.; Kamiya K.;
RA Karasawa W.; Katagiri S.; Kikuta A.; Kobayashi N.; Kono I.;
RA Machita K.; Maehara T.; Mizuno H.; Mizubayashi T.; Mukai Y.;
RA Nagasaki H.; Nakashima M.; Nakama Y.; Nakamichi Y.; Nakamura M.;
RA Namiki N.; Negishi M.; Ohta I.; Ono N.; Saji S.; Sakai K.; Shibata M.;
RA Shimokawa T.; Shomura A.; Song J.; Takazaki Y.; Terasawa K.; Tsuji K.;
RA Waki K.; Yamagata H.; Yamane H.; Yoshiki S.; Yoshihara R.; Yukawa K.;
RA Zhong H.; Iwama H.; Endo T.; Ito H.; Hahn J.H.; Kim H.-I.; Eun M.-Y.;
RA Yano M.; Jiang J.; Gojobori T.;
RT "The genome sequence and structure of rice chromosome 1.";
RL Nature 420:312-316(2002).
DR EMBL; AP003380; BAB92567.1; -; Genomic_DNA.
DR Gramene; Q8LIW2; -.
SQ SEQUENCE 300 AA; 32009 MW; D4B70188E3541B3A CRC64;

Query Match 100.0%; Score 45; DB 2; Length 300;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 191 RPQPPPP 197

RESULT 11
Q6MP85 BDEBA
ID Q6MP85 BDEBA PRELIMINARY; PRT; 317 AA.
AC Q6MP85;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative hydrolase precursor.
```

GN OrderedLocusNames=Bd0978;
OS Bdellovibrio bacteriovorus.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Bdellovibrionales;
OC Bdellovibrionaceae; Bdellovibrio.
OX NCBI_TaxID=959;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HD100 / DSM 50701 / ATCC 15356 / NCIB 9529;
RX PubMed=14752164; DOI=10.1126/science.1093027;
RA Rendulic S., Jagtap P., Rosinus A., Eppinger M., Baar C., Lanz C.,
RA Keller H., Lambert C., Evans K.J., Goemann A., Meyer F.,
RA Sockett R.E., Schuster S.C.;
RT "A predator unmasked: life cycle of Bdellovibrio bacteriovorus from a
RT genomic perspective.";
RL Science 303:689-692(2004).
DR EMBL; BX842648; CAE78913.1; -; Genomic_DNA.
KW Complete proteome; Hydrolase; Signal.
FT SIGNAL 1 8 Potential.
SQ SEQUENCE 317 AA; 35558 MW; CC4BB5707945FD36 CRC64;

Query Match 100.0%; Score 45; DB 2; Length 317;
Best Local Similarity 100.0%; Pred. NO. 3.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 300 RPQPPPP 306
|||

RESULT 12
Q4T9Z6_TETNG
ID Q4T9Z6_TETNG PRELIMINARY; PRT; 328 AA.
AC Q4T9Z6;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome undetermined SCAF7480, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG0004514001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAEE01007480; CAF90286.1; -; Genomic_DNA.
FT NON TER 1 1
SQ SEQUENCE 328 AA; 34890 MW; 4020AD04402C59BE CRC64;

Query Match 100.0%; Score 45; DB 2; Length 328;
Best Local Similarity 100.0%; Pred. NO. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 179 RPQPPPP 185
|||

RESULT 13
Q9LHG8_ARATH
ID Q9LHG8_ARATH PRELIMINARY; PRT; 398 AA.
AC Q9LHG8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE And human tumor susceptibility gene-like protein (Hypothetical protein
DE At3g12400) (Hypothetical protein T2E22.28).
GN Name=At3g12400; Synonyms=T2E22.28;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; euroids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20363099; PubMed=10907853;
RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the 4,251,695 bp regions covered by 90 P1, TAC
RT and BAC clones.";
RL DNA Res. 7:217-221(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22088475; PubMed=12093376;
RA Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
RT annotation.";
RL Genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE.
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
RA Maiti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R.,
RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [7]
RP NUCLEOTIDE SEQUENCE.
RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hsuan V.W.,
RA Lee J.M., Quach H.L., Tang C., Toriumi M., Wallender E.K., Wong C.,
RA Wu H.C., Yu G., Yuan S., Bowser L., Carninci P., Chen H., Cheuk R.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Kim C., Lam B., Lin J., Miranda M., Narusaka M., Nguyen M.,
RA Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A.,

RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP002047; BAB03147.1; -; Genomic_DNA.
DR EMBL; AY056283; AAL07132.1; -; mRNA.
DR EMBL; AY087782; AAM65318.1; -; mRNA.
DR EMBL; AC069474; AAG51025.1; -; Genomic_DNA.
DR EMBL; AY133779; AAM91713.1; -; mRNA.
DR HSSP; Q99816; 1KPQ.
DR GO; GO:0004840; F:ubiquitin conjugating enzyme activity; IEA.
DR GO; GO:0015031; P:protein transport; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR008883; Tsgl01.
DR InterPro; IPR000608; UBQ-conjugat_E2.
DR Pfam; PF05743; Tsgl01; 1.
DR SMART; SM00212; UBCC; 1.
KW Hypothetical protein.
SQ SEQUENCE 398 AA; 44715 MW; 659996F9784FDC20 CRC64;

Query Match 100.0%; Score 45; DB 2; Length 398;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
|||
Db 163 RPQPPPP 169

RESULT 14
Q9C2D6 NEUCR
ID Q9C2D6_NEUCR PRELIMINARY; PRT; 416 AA.
AC Q9C2D6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein 9G6.290.
GN Name=9G6.290;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Schulte U., Algn V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA German Neurospora genome project;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Catalyzes the methyl transfer from S-adenosyl-methionine
CC to the C-24 of zymosterol to form fecosterol (By similarity).
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 5-alpha-cholesta-
CC 8,24-dien-3-beta-ol = S-adenosyl-L-homocysteine + 24-methylene-5-
CC alpha-cholest-8-en-3-beta-ol.
CC -!- PATHWAY: Ergosterol biosynthesis.
DR EMBL; AL513463; CAC28777.2; -; Genomic_DNA.
DR GO; GO:0008825; F:cyclopropane-fatty-acyl-phospholipid syntha. . .; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0016126; P:sterol biosynthesis; IEA.
DR InterPro; IPR003333; CMAS.
DR InterPro; IPR001601; Methyltransf.
DR InterPro; IPR000051; SAM_bind.
DR Pfam; PF02353; CMAS; 1.
KW Hypothetical protein; Methyltransferase; Steroid biosynthesis;
KW Sterol biosynthesis; Transferase.
SQ SEQUENCE 416 AA; 46037 MW; 9A55846043C17B2F CRC64;

Query Match 100.0%; Score 45; DB 2; Length 416;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
|||
.

Db 349 RPQPPPP 355
RESULT 15
Q9VZC2 DROME
ID Q9VZC2_DROME PRELIMINARY; PRT; 420 AA.
AC Q9VZC2; Q8SZ47;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE CG15021-PA (REI7165p).
GN Name=CG15021; ORFNames=CG15021;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadeniou E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Murphy L., Muzny D.M., Nelson D.L.,
RA Mount S.M., Moy M., Murphy B., Murphy C., Morris J., Moshrefi A.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
RT melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426070; PubMed=12537573;


```
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirkas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
a Genomics perspective.";
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RG Berkeley Drosophila Genome Project;
RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirkas R., Smith E.,
RA Yu C., Rubin G.;
RT "Drosophila melanogaster release 4 sequence.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE.
RG FlyBase;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
RN [7]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003481; AAF47902.2; -; Genomic_DNA.
DR EMBL; AY071124; AAL48746.1; -; mRNA.
DR Ensembl; CG15021; Drosophila melanogaster.
DR FlyBase; FBgn0035544; CG15021.
DR GO; GO:0005199; F:structural constituent of cell wall; IEA.
DR InterPro; IPR003882; Pistil_extensin.
DR InterPro; IPR004019; YLP_motif.
DR Pfam; PF02757; YLP; 2.
DR PRINTS; PR01218; PSTIEXTENSIN.
SQ SEQUENCE 420 AA; 42947 MW; 96D62FFDC9F996E1 CRC64;

Query Match 100.0%; Score 45; DB 2; Length 420;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 205 RPQPPPP 211

RESULT 16
Q9PZ18 GVXN PRELIMINARY; PRT; 449 AA.
AC Q9PZ18;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE ORF25.
GN Name=ORF25;
OS Xestia c-nigrum granulosis virus (XnGV) (Xestia c-nigrum
OS granulovirus).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae; Granulovirus.
```

```
OX NCBI_TaxID=51677;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=99434230; PubMed=10502508; DOI=10.1006/viro.1999.9894;
RA Hayakawa T., Ko R., Okano K., Seong S.I., Goto C., Maeda S.;
RT "Sequence analysis of the Xestia c-nigrum granulovirus genome.";
RL Virology 262:277-297(1999).
DR EMBL; AF162221; AAF05139.1; -; Genomic_DNA.
SQ SEQUENCE 449 AA; 51281 MW; 7B7F5666AECABF23 CRC64;

Query Match 100.0%; Score 45; DB 2; Length 449;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 332 RPQPPPP 338

RESULT 17
Q4POX1 USTMA PRELIMINARY; PRT; 551 AA.
ID Q4POX1_USTMA PRELIMINARY;
AC Q4POX1;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Predicted protein.
GN ORFNames=UM06242.1;
OS Ustilago maydis 521.
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OX NCBI_TaxID=237631;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=521;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
RA Bayul T., Blitshsteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collymore A., Considine T., Cook K., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA Gnirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamvysselis M., Karlsson E.,
RA Kells C., Kieu A., Kisner P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., Mcghee T., Meldrim J., Meneus L.,
RA Mesirov J., Mihalev A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., News C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA O'Neill K., Osman S., Parker S., Perrin D., Phunkhang P., Pigani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thoulutsang Y., Topham K.,
RA Towey S., Tsamla T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
```

RT "The genome sequence of Ustilago maydis.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACP01000237; EAK87122.1; -; Genomic DNA.
SQ SEQUENCE 551 AA; 61349 MW; 8928BC5EE48FC91 CRC64;

Query Match 100.0%; Score 45; DB 2; Length 551;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
Db 198 RPQPPPP 204

RESULT 18
Q525TS MAGGR
ID Q525TS_MAGGR PRELIMINARY; PRT; 587 AA.
AC Q525TS;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=MG06971.4;
OS Magnaporthe grisea 70-15.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.
OX NCBI_TaxID=242507;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
RA Bayul T., Blitshsteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collimore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA Gnrirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatchner B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kanat A., Kamvyselis M., Karlsson E.,
RA Kells C., Kieu A., Kisner P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., Mcghee T., Meldrim J., Meneus L.,
RA Mesirov J., Mihalev A., Mihova T., Mikkelsen T., Mienga V., Moru K.,
RA Mozes J., Munson G., Munson G., Naylor J., Neues C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
RA O'neill K., Osman S., Parker S., Perrin D., Phunkhang P., Piquani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzey P., Testaye S., Theodore J., Thoulutsang Y., Topham K.,
RA Towey S., Tsamla T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Magnaporthe grisea.";
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.

[2]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=70-15;
RC Dean R., Mitchell T., Brown D., Pan H., Thon M.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Zhu H., Blackmon B.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACU01000620; EAA55314.1; -; Genomic DNA.
DR InterPro; IPR000232; HSF_DNA_bd.
DR InterPro; IPR002341; HSF_ETS_DNA_bd.
DR InterPro; IPR011991; Wing_hlx_DNA_bd.
DR Pfam; PF00447; HSF_DNA-bind; 1.
DR ProDom; PD001788; HSF_DNA_bind; 1.
DR SMART; SM00415; HSF; 1.
KW DNA-binding; Hypothetical protein; Nuclear protein; Transcription.
SQ SEQUENCE 587 AA; 64390 MW; 11B6F08B03A3F936 CRC64;

Query Match 100.0%; Score 45; DB 2; Length 587;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPQPPPP 7
Db 409 RPQPPPP 415

RESULT 19
Q4TA45 TETNG
ID Q4TA45_TETNG PRELIMINARY; PRT; 624 AA.
AC Q4TA45;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 10 SCAF7449, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00004438001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAE01007449; CAF90237.1; -; Genomic DNA.

```
FT  NON TER      624      624
SQ  SEQUENCE     624 AA;  69882 MW;  E186BFB8766BA752 CRC64;

Query Match      100.0%;  Score 45;  DB 2;  Length 624;
Best Local Similarity 100.0%;  Pred. No. 6.9e+02;
Matches 7;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

QY  1 RPQPPPP 7
    |||||||
Db  57 RPQPPPP 63

RESULT 20
Q5H9B0 HUMAN
ID  Q5H9B0_HUMAN PRELIMINARY;      PRT;    669 AA.
AC  Q5H9B0;
DT  10-MAY-2005 (TrEMBLrel. 30, Created)
DT  10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT  13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE  Synapsin I.
GN  Name=SYN1; ORFNames=RP1-230G1.2-002;
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC  Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RA  Wray P.;
RL  Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN  [2]
RP  NUCLEOTIDE SEQUENCE.
RA  Grafham D.;
RL  Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR  EMBL; Z84466; CAI42461.1; -; Genomic DNA.
DR  EMBL; AL009172; CAI42445.1; -; Genomic DNA.
DR  EMBL; AL009172; CAI42461.1; JOINED; Genomic DNA.
DR  EMBL; Z84466; CAI42445.1; JOINED; Genomic DNA.
DR  SMR; Q5H9B0; 112-417.
DR  GO; GO:0008021; C:synaptic vesicle; IEA.
DR  GO; GO:0007269; P:neurotransmitter secretion; IEA.
DR  InterPro; IPR001359; Synapsin.
DR  Pfam; PF02750; Synapsin_C; 1.
DR  Pfam; PF02078; Synapsin_N; 1.
DR  PRINTS; PR01368; SYNAPSIN.
DR  PROSITE; PS00415; SYNAPSIN_1; 1.
DR  PROSITE; PS00416; SYNAPSIN_2; 1.
SQ  SEQUENCE 669 AA; 70033 MW;  5E400115415D3E32 CRC64;

Query Match      100.0%;  Score 45;  DB 2;  Length 669;
Best Local Similarity 100.0%;  Pred. No. 7.4e+02;
Matches 7;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

QY  1 RPQPPPP 7
    |||||||
Db  27 RPQPPPP 33

RESULT 21
Q4G0U1_HUMAN
ID  Q4G0U1_HUMAN PRELIMINARY;      PRT;    687 AA.
AC  Q4G0U1;
DT  13-SEP-2005 (TrEMBLrel. 31, Created)
DT  13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT  13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE  Hypothetical protein (fragment).
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC  Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
```

```
RC  TISSUE=Brain;
RX  MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA  Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA  Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA  Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA  Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA  Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA  Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA  Brownstein M.J., Udín T.B., Toshiyuki S., Carninci P., Prange C.,
RA  Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA  Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA  Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA  Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA  Fahey J., Helton E., Kettman M., Madan A., Shevchenko Y., Bouffard G.G.,
RA  Whiting M., Madan A., Young A.C., Shevchenko Y., Smailus D.E.,
RA  Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA  Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA  Butterfield Y.S.N., Krzywinski M.I., Skalska U., Marra M.A.,
RA  Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT  "Generation and initial analysis of more than 15,000 full-length human
RT  and mouse cDNA sequences.";
RL  Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN  [2]
RP  NUCLEOTIDE SEQUENCE.
RC  TISSUE=Brain;
RG  NIH MGC Project;
RL  Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR  EMBL; BC036690; AAH36690.1; -; mRNA.
KW  Hypothetical protein.
FT  NON TER      1
SQ  SEQUENCE 687 AA; 71080 MW;  E5EA157E801281CD CRC64;

Query Match      100.0%;  Score 45;  DB 2;  Length 687;
Best Local Similarity 100.0%;  Pred. No. 7.6e+02;
Matches 7;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

QY  1 RPQPPPP 7
    |||||||
Db  369 RPQPPPP 375

RESULT 22
SYN1_RAT
ID  SYN1_RAT      STANDARD;      PRT;    704 AA.
AC  P09951; Q9WUX7;
DT  01-MAR-1989 (Rel. 10, Created)
DT  30-MAY-2000 (Rel. 39, Last sequence update)
DT  13-SEP-2005 (Rel. 48, Last annotation update)
DE  Synapsin-1 (Synapsin I).
GN  Name=Syn1;
OS  Rattus norvegicus (Rat).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Euthera; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC  Muridea; Muridae; Murinae; Rattus.
OX  NCBI_TaxID=10116;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RC  TISSUE=Brain;
RX  MEDLINE=89388265; PubMed=25066642;
RA  Suedhof T.C., Czernik A.J., Kao H.-T., Takei K., Johnston P.A.,
RA  Horiuchi A., Kanazir S.D., Wagner M.A., Perin M.S., de Camilli P.,
RA  Greengard P.;
RT  "Synapsins: mosaics of shared and individual domains in a family of
RT  synaptic vesicle phosphoproteins.";
RL  Science 245:1474-1480(1989).
RN  [2]
RP  NUCLEOTIDE SEQUENCE.
RC  TISSUE=Brain;
RX  MEDLINE=87133474; PubMed=30287773;
RA  McCaffery C.A., Degennaro L.J.;
RT  "Determination and analysis of the primary structure of the nerve
RT  terminal specific phosphoprotein, synapsin I.";
RL  EMBO J. 5:3167-3173(1986).
```


RN RP CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=99313567; PubMed=10386995;
RA Cole R.N., Hart G.W.;
RT "Glycosylation sites flank phosphorylation sites on synapsin I: O-linked N-acetylglucosamine residues are localized within domains mediating synapsin I interactions.";
RL J. Neurochem. 73:418-428(1999).
[4]
RN RP CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=22326970; PubMed=12438562; DOI=10.1074/mcp.M200048-MCP200; Wells L., Vosseller K., Cole R.N., Cronshaw J.M., Matunis M.J., Hart G.W.;
RT "Mapping sites of O-GlcNAc modification using affinity tags for serine and threonine post-translational modifications.";
RL Mol. Cell. Proteomics 1:791-804(2002).
[5]
RN RP INTERACTIONS WITH NOS1 AND CAPON.
RX MEDLINE=21874142; PubMed=11867766; DOI=10.1073/pnas.261705799; Jaffrey S.R., Benfenati F., Snowman A.M., Czernik A.J., Snyder S.H.;
RT "Neuronal nitric-oxide synthase localization mediated by a ternary complex with synapsin and CAPON.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:3199-3204(2002).
CC -!- FUNCTION: Neuronal phosphoprotein that coats synaptic vesicles, binds to the cytoskeleton, and is believed to function in the regulation of neurotransmitter release.
CC -!- SUBUNIT: Homodimer (By similarity). Interacts with CAPON. Forms a ternary complex with NOS1.
CC -!- SUBCELLULAR LOCATION: Synapse.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=IA;
CC IsoId=P09951-1; Sequence=Displayed;
CC Name=IB;
CC IsoId=P09951-2; Sequence=VSP_006318, VSP_006319;
CC -!- PTM: Substrate of at least four different protein kinases. It is probable that phosphorylation plays a role in the regulation of synapsin-1 in the nerve terminal (By similarity).
CC -!- SIMILARITY: Belongs to the synapsin family.
CC -!- CAUTION: Ref.2 sequence was incorrect at many places due to sequencing errors.

This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

DR EMBL; M27812; AAA42145.1; -; mRNA.
DR EMBL; M27924; AAA42148.1; -; mRNA.
DR EMBL; X04655; CAA28353.1; ALT_SEQ; mRNA.
DR PIR; A25704; A25704.
DR PIR; A30411; A30411.
DR PDB; 1PK8; X-ray; A/B/C/D/E/F/G/H=2-421.
DR PDB; 1PX2; X-ray; A/B=2-421.
DR Ensembl; ENSRNOG0000010365; Rattus norvegicus.
DR RGD; 3797; SynI.
DR GO; GO:0008021; C:synaptic vesicle; IDA.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0007269; P:neurotransmitter secretion; TAS.
DR InterPro; IPR001359; Synapsin.
DR PANTHER; PTHR10841; Synapsin; 1.
DR Pfam; PF02750; Synapsin_C; 1.
DR Pfam; PF02078; Synapsin_N; 1.
DR PRINTS; PR01368; SYNAPSIN.
DR PROSITE; PS00415; SYNAPSIN_1; 1.
DR PROSITE; PS00416; SYNAPSIN_2; 1.
DR 3D-structure; Actin-binding; Alternative splicing; Glycoprotein; Phosphorylation; Repeat; Synapse.
FT REGION 1 28 A.
FT REGION 29 112 B; linker.
FT REGION 113 420 C; actin-binding and synaptic-vesicle binding.

FT REGION 421 655 D; Pro-rich linker.
FT REGION 656 704 E.
FT MOD_RES 9 Phosphoserine (by CamK1 and PKA).
FT MOD_RES 566 566 Phosphoserine (by CamK2).
FT MOD_RES 603 603 Phosphoserine (by CamK2).
FT CARBOHYD 55 O-linked (GlcNAc. .).
FT CARBOHYD 56 56 O-linked (GlcNAc. .).
FT CARBOHYD 87 87 O-linked (GlcNAc. .).
FT CARBOHYD 96 96 O-linked (GlcNAc. .).
FT CARBOHYD 103 103 O-linked (GlcNAc. .).
FT CARBOHYD 261 261 O-linked (GlcNAc. .).
FT CARBOHYD 430 430 O-linked (GlcNAc. .).
FT CARBOHYD 516 516 O-linked (GlcNAc. .).
FT CARBOHYD 524 524 O-linked (GlcNAc. .).
FT CARBOHYD 562 562 O-linked (GlcNAc. .).
FT CARBOHYD 576 576 O-linked (GlcNAc. .).
FT VARSPLIC 660 668 NKSQLTNA -> KASPAQAQP (in isoform IB). /FTId=VSP_006318.
FT VARSPLIC 669 704 Missing (in isoform IB). /FTId=VSP_006319.
FT SEQUENCE 704 AA; 73988 MW; 65799FE7CFE18B5 CRC64;
Query Match 100.0%; Score 45; DB 1; Length 704;
Best Local Similarity 100.0%; Pred. No. 7.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RPQPPPP 7
Db 27 RPQPPPP 33
RESULT 23
Q95XQ7 CAEEL
ID Q95XQ7 CAEEL PRELIMINARY; PRT; 704 AA.
AC Q95XQ7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN ORFNames=Y39G10AR.15;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; AC025716; AAK39604.1; -; Genomic DNA.
DR Ensembl; Y39G10AR.15; Caenorhabditis elegans.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 704 AA; 80412 MW; A2B949D866777EF8 CRC64;
Query Match 100.0%; Score 45; DB 2; Length 704;
Best Local Similarity 100.0%; Pred. No. 7.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RPQPPPP 7
Db 143 RPQPPPP 149
RESULT 24
SYN1 HUMAN
ID SYN1_HUMAN STANDARD; PRT; 705 AA.
AC P17600; O75825; Q5H9A9;
DT 01-AUG-1990 (Rel. 15, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)

DE Synapsin-1 (Synapsin I) (Brain protein 4.1).
GN Name=SYN1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA], AND ALTERNATIVE SPLICING.
RC TISSUE=Brain;
RX MEDLINE=90243651; PubMed=2110562;
RA Suedhof T.C.;
RT "The structure of the human synapsin I gene and protein.";
RL J. Biol. Chem. 265:7849-7852(1990).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15772651; DOI=10.1038/nature03440;
RA Ross M.T., Grafham D.V., Coffey A.J., Scherer S., McLay K., Muzny D.,
RA Platzer M., Howell G.R., Burrows C., Bird C.P., Frankish A.,
RA Lovell F.L., Howe K.L., Ashurst J.L., Fulton R.S., Sudbrak R., Wen G.,
RA Jones M.C., Hurler M.E., Andrews T.D., Scott C.E., Searle S.,
RA Ramser J., Whittaker A., Deadman R., Carter N.P., Hunt S.E., Chen R.,
RA Cree A., Gunaratne P., Havlak P., Hodgson A., Metzker M.L.,
RA Richards S., Scott G., Steffen D., Ambrose K.D., Ansari-Lari M.A., Aradhya S.,
RA Worley K.C., Ainscough R., Ambrose K.D., Sodergren E., Wheeler D.A.,
RA Ashwell R.I., Babbage A.K., Bagguley C.L., Ballabio A., Banerjee R.,
RA Barker G.E., Barlow K.F., Barrett I.P., Bates K.N., Beare D.M.,
RA Beasley H., Beasley O., Beck A., Bethel G., Blechschmidt K., Brady N.,
RA Bray-Allen S., Bridgeman A.M., Brown A.J., Brown M.J., Bonnin D.,
RA Bruford E.A., Buhay C., Burch P., Burford D., Burgess J., Burrill W.,
RA Burton J., Bye J.M., Carder C., Carrel L., Chako J., Chapman J.C.,
RA Chavez D., Chen E., Chen G., Chen Y., Chen Z., Chinault C.,
RA Ciccodicola A., Clark S.Y., Clarke G., Clee C.M., Clegg S.,
RA Clerc-Blankenburg K., Clifford K., Cobley V., Cole C.G., Conquer J.S.,
RA Corby N., Connor R.E., David R., Davies J., Davis C., Davis J.,
RA Delgado O., Deshazo D., Dhami P., Ding Y., Dinh H., Dodsworth S.,
RA Draper H., Dugan-Rocha S., Dunham A., Dunn M., Durbin K.J., Dutta I.,
RA Eades T., Ellwood M., Emery-Cohen A., Errington H., Evans K.L.,
RA Faulkner L., Francis F., Frankland J., Fraser A.E., Gargocsy P.,
RA Gilbert J., Gill R., Gloeckner G., Gregory S.G., Gribble S.,
RA Griffiths C., Grocock R., Gu Y., Gwilliam R., Hamilton C., Hart E.A.,
RA Hawes A., Heath P.D., Heitmann K., Hennig S., Hernandez J.,
RA Hinzmann B., Ho S., Hoffs M., Howden P.J., Huckle E.J., Hume J.,
RA Hunt P.J., Hunt A.R., Isherwood J., Jacob L., Johnson D., Jones S.,
RA de Jong P.J., Joseph S.S., Keenan S., Kelly S., Kershaw J.K., Khan Z.,
RA Kioschis P., Klages S., Knights A.J., Kosiura A., Kovar-Smith C.,
RA Laird G.K., Langford C., Lawlor S., Leversha M., Lewis L., Liu W.,
RA Lloyd C., Lloyd D.M., Loulseghe H., Loveland J.E., Lovell J.D.,
RA Lozado R., Lu J., Lyne R., Ma J., Maheshwari M., Matthews L.H.,
RA McDowall J., McLaren S., McMurray A., Meidl P., Meitinger T.,
RA Milne S., Miner G., Mistry S.L., Morgan M., Morris S., Mueller I.,
RA Mullikin J.C., Nguyen N., Nordsiek G., Nyakatura G., O'dell C.N.,
RA Okwuonu G., Palmer S., Pandian R., Parker D., Parrish J.,
RA Pasternak S., Patel D., Pearce A.V., Pearson D.M., Pelan S.E.,
RA Perez L., Porter K.M., Ramsey Y., Reichwald K., Rhodes S.,
RA Ridler K.A., Schlessinger D., Schueler M.G., Sehra H.K.,
RA Shaw-Smith C., Shen H., Sheridan E.M., Shownkeen R., Skuce C.D.,
RA Smith M.L., Sotharan E.C., Steingruber H.E., Steward C.A., Storey R.,
RA Swann R.M., Swarbreck D., Tabor P.E., Taudien S., Taylor T.,
RA Teague B., Thomas K., Thorpe A., Timms K., Tracey A., Trevanion S.,
RA Tromans A.C., d'Urso M., Verduzco D., Villasana D., Waldron L.,
RA Wall M., Wang Q., Warren J., Warry G.L., Wei X., West A.,
RA Whitehead S.L., Whiteley M.N., Wilkinson J.E., Willey D.L.,
RA Williams G., Williams L., Williamson A., Williamson H., Wilming L.,
RA Woodmansey R.L., Wray P.W., Yen J., Zhang J., Zhou J., Zoghbi H.,
RA Zorilla S., Buck D., Reinhardt R., Pouster A., Rosenthal A.,
RA Lehrach H., Meindl A., Minx P.J., Hillier L.W., Willard H.F.,
RA Wilson R.K., Waterston R.H., Rice C.M., Vaudin M., Coulson A.,
RA Nelson D.L., Weinstock G., Sulston J.E., Durbin R., Hubbard T.,
RA Gibbs R.A., Beck S., Rogers J., Bentley D.R.;
RT "The DNA sequence of the human X chromosome.";
RL Nature 434:325-337(2005).
RN [3]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 1-125.
RX MEDLINE=90368667; PubMed=2118519;
RA Sauerwald A., Hoesche C., Oschwald R., Kilmann M.W.;
RT "The 5'-flanking region of the synapsin I gene. A G+C-rich, TATA- and
RT CAAT-less, phylogenetically conserved sequence with cell type-specific
RT promoter function.";
RL J. Biol. Chem. 265:14932-14937(1990).
RN [4]
RP INVOLVEMENT IN EPILEPSY.
RX PubMed=14985377; DOI=10.1136/jmg.2003.013680;
RA Garcia C.C., Blair H.J., Seager M., Coulthard A., Tennant S.,
RA Buddles M., Curtis A., Goodship J.A.;
RT "Identification of a mutation in synapsin I, a synaptic vesicle
RT protein, in a family with epilepsy.";
RL J. Med. Genet. 41:183-187(2004).
CC -!- FUNCTION: Neuronal phosphoprotein that coats synaptic vesicles,
CC binds to the cytoskeleton, and is believed to function in the
CC regulation of neurotransmitter release. The complex formed with
CC NOS1 and CAPON proteins is necessary for specific nitric-oxid
CC functions at a presynaptic level.
CC -!- SUBUNIT: Homodimer. Interacts with CAPON. Forms a ternary complex
CC with NOS1 (By similarity).
CC -!- SUBCELLULAR LOCATION: Synapse.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=IA;
CC IsoId=P17600-1; Sequence=Displayed;
CC Name=IB;
CC IsoId=P17600-2; Sequence=VSP_006316, VSP_006317;
CC -!- PTM: Substrate of at least four different protein kinases. It is
CC probable that phosphorylation plays a role in the regulation of
CC synapsin-1 in the nerve terminal.
CC -!- DISEASE: Defects in SYN1 are a cause of X-linked recessive
CC epilepsy associated with variable learning disabilities and
CC behavior disorders [MIM:300491].
CC -!- SIMILARITY: Belongs to the synapsin family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; M58378; AAC41930.1; -; Genomic DNA.
DR EMBL; M58321; AAC41930.1; JOINED; Genomic DNA.
DR EMBL; M58341; AAC41930.1; JOINED; Genomic DNA.
DR EMBL; M58351; AAC41930.1; JOINED; Genomic DNA.
DR EMBL; M58353; AAC41930.1; JOINED; Genomic DNA.
DR EMBL; M58359; AAC41930.1; JOINED; Genomic DNA.
DR EMBL; M58371; AAC41930.1; JOINED; Genomic DNA.
DR EMBL; M58372; AAC41930.1; JOINED; Genomic DNA.
DR EMBL; M58373; AAC41930.1; JOINED; Genomic DNA.
DR EMBL; M58374; AAC41930.1; JOINED; Genomic DNA.
DR EMBL; M58375; AAC41930.1; JOINED; Genomic DNA.
DR EMBL; M58376; AAC41930.1; JOINED; Genomic DNA.
DR EMBL; M58377; AAC41930.1; JOINED; Genomic DNA.
DR EMBL; M58378; AAC41931.1; ALT SEQ; Genomic DNA.
DR EMBL; M58321; AAC41931.1; JOINED; Genomic DNA.
DR EMBL; M58341; AAC41931.1; JOINED; Genomic DNA.
DR EMBL; M58351; AAC41931.1; JOINED; Genomic DNA.
DR EMBL; M58353; AAC41931.1; JOINED; Genomic DNA.
DR EMBL; M58359; AAC41931.1; JOINED; Genomic DNA.
DR EMBL; M58371; AAC41931.1; JOINED; Genomic DNA.
DR EMBL; M58372; AAC41931.1; JOINED; Genomic DNA.
DR EMBL; M58373; AAC41931.1; JOINED; Genomic DNA.
DR EMBL; M58374; AAC41931.1; JOINED; Genomic DNA.
DR EMBL; M58375; AAC41931.1; JOINED; Genomic DNA.
DR EMBL; M58376; AAC41931.1; JOINED; Genomic DNA.
DR EMBL; M58377; AAC41931.1; JOINED; Genomic DNA.
DR EMBL; AL009172; CAI42446.1; -; Genomic DNA.
DR EMBL; Z84466; CAI42446.1; JOINED; Genomic DNA.
DR EMBL; Z84466; CAI42462.1; -; Genomic DNA.
DR EMBL; AL009172; CAI42462.1; JOINED; Genomic DNA.

DR EMBL; M55301; AAA60608.1; -; Genomic_DNA.
DR PIR; A35363; A35363.
DR HSSP; P17599; 1AUX.
DR SMR; P17600; 112-417.
DR Ensembl; ENSG00000008056; Homo sapiens.
DR HGNC; HGNC:11494; SYN1.
DR MIM; 313440; -.
DR MIM; 300491; -.
DR GO; GO:0005215; F:transporter activity; TAS.
DR GO; GO:0007268; P:synaptic transmission; TAS.
DR InterPro; IPR001359; Synapsin.
DR PANTHER; PTHR10841; Synapsin; 1.
DR Pfam; PF02750; Synapsin_C; 1.
DR Pfam; PF02078; Synapsin_N; 1.
DR PRINTS; PR01368; SYNAPSIN.
DR PROSITE; PS00415; SYNAPSIN_1; 1.
DR PROSITE; PS00416; SYNAPSIN_2; 1.
KW Actin-binding; Alternative splicing; Epilepsy; Phosphorylation;
KW Repeat; Synapse.
FT REGION 1 28 A.
FT REGION 29 112 B; linker.
FT REGION 113 420 C; actin-binding and synaptic-vesicle
FT binding.
FT REGION 421 655 D; Pro-rich linker.
FT REGION 656 705 E.
FT MOD_RES 9 Phosphoserine (by CaMK1 and PKA).
FT MOD_RES 568 Phosphoserine (by CaMK2).
FT MOD_RES 605 Phosphoserine (by CaMK2).
FT VARSPLIC 661 669 NKQSLTNA -> KASPAQAQP (in isoform IB).
FT /FTid=VSP_006316.
FT VARSPLIC 670 705 Missing (in isoform IB).
FT /FTid=VSP_006317.
FT CONFLICT 138 138 E -> G (in Ref. 1).
FT CONFLICT 631 631 R -> A (in Ref. 1).
SQ SEQUENCE 705 AA; 74111 MW; BE4CE46C942300B0 CRC64;

Query Match 100.0%; Score 45; DB 1; Length 705;
Best Local Similarity 100.0%; Pred. No. 7.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 27 RPQPPPP 33

RESULT 25
SYN1_BOVIN STANDARD; PRT; 706 AA.
AC P17599;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Synapsin-1 (Synapsin I).
GN Name=SYN1;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1] NUCLEOTIDE SEQUENCE [MRNA] (ISOFORMS IA AND IB).
RP TISSUE=Brain;
RC MEDLINE=89388265; PubMed=2506642;
RA Suedhof T.C., Czernik A.J., Kao H.-T., Takei K., Johnston P.A.,
RA Horiuchi A., Kanazir S.D., Wagner M.A., Perin M.S., de Camilli P.,
RA Greengard P.;
RT "Synapsins: mosaics of shared and individual domains in a family of
RT synaptic vesicle phosphoproteins.";
RL Science 245:1474-1480(1989).
RN [2]
RP PHOSPHORYLATION SITE BY PROLINE-DIRECTED KINASE.
RX MEDLINE=90216728; PubMed=2108963;
RA Hall F.L., Mitchell J.P., Vulliamt P.R.;

RT "Phosphorylation of synapsin I at a novel site by proline-directed
RT protein kinase.";
RL J. Biol. Chem. 265:6944-6948(1990).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.15 ANGSTROMS) OF 112-417.
RX MEDLINE=98130589; PubMed=9463376; DOI=10.1093/emboj/17.4.977;
RA Esser L., Wang C.-R., Hosaka M., Smagula C.S., Suedhof T.C.,
RA Deisenhofer J.;
RT "Synapsin I is structurally similar to ATP-utilizing enzymes.";
RL EMBO J. 17:977-984(1998).
CC -!- FUNCTION: Neuronal phosphoprotein that coats synaptic vesicles,
CC binds to the cytoskeleton, and is believed to function in the
CC regulation of neurotransmitter release. The complex formed with
CC NOS1 and CAPON proteins is necessary for specific nitric-oxid
CC functions at a presynaptic level (By similarity).
CC -!- SUBUNIT: Homodimer. Interacts with CAPON. Forms a ternary complex
CC with NOS1 (By similarity).
CC -!- SUBCELLULAR LOCATION: Synapse.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=IA;
CC IsoId=P17599-1; Sequence=Displayed;
CC Name=IB;
CC IsoId=P17599-2; Sequence=VSP_006314, VSP_006315;
CC -!- PTM: Substrate of at least four different protein kinases.
CC Phosphorylation may play a role in the regulation of synapsin-1 in
CC the nerve terminal.
CC -!- SIMILARITY: Belongs to the synapsin family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; M27810; AAA30761.1; -; mRNA.
DR EMBL; M27811; AAA30762.1; -; mRNA.
DR PDB; 1AUV; X-ray; A/B=110-420.
DR PDB; 1AUX; X-ray; A/B=110-420.
DR GlycoSuiteDB; P17599; -.
DR InterPro; IPR001359; Synapsin.
DR PANTHER; PTHR10841; Synapsin; 1.
DR Pfam; PF02750; Synapsin_C; 1.
DR Pfam; PF02078; Synapsin_N; 1.
DR PRINTS; PR01368; SYNAPSIN.
DR PROSITE; PS00415; SYNAPSIN_1; 1.
DR PROSITE; PS00416; SYNAPSIN_2; 1.
KW 3D-structure; Actin-binding; Alternative splicing; Phosphorylation;
KW Repeat; Synapse.
FT REGION 1 28 A.
FT REGION 29 112 B; linker.
FT REGION 113 420 C; actin-binding and synaptic-vesicle
FT binding.
FT REGION 421 656 D; Pro-rich linker.
FT REGION 657 706 E.
FT MOD_RES 9 Phosphoserine (by CaMK1 and PKA).
FT MOD_RES 551 Phosphoserine (by PDPK).
FT MOD_RES 568 Phosphoserine (by CaMK2).
FT MOD_RES 605 Phosphoserine (by CaMK2).
FT VARSPLIC 662 670 NKQSLTNA -> KASPAQAQP (in isoform IB).
FT /FTid=VSP_006314.
FT VARSPLIC 671 706 Missing (in isoform IB).
FT /FTid=VSP_006315.
FT STRAND 114 119
FT TURN 122 123
FT HELIX 126 130
FT TURN 131 132
FT STRAND 134 135
FT TURN 136 138
FT STRAND 139 146
FT TURN 149 150
FT STRAND 151 155
FT TURN 157 158

FT	STRAND	161	165		
FT	TURN	172	173		
FT	STRAND	176	178		
FT	STRAND	182	185		
FT	STRAND	190	190		
FT	TURN	194	195		
FT	STRAND	197	197		
FT	HELIX	199	207		
FT	TURN	208	209		
FT	STRAND	212	213		
FT	HELIX	216	221		
FT	TURN	222	223		
FT	HELIX	225	239		
FT	TURN	241	243		
FT	STRAND	246	246		
FT	STRAND	250	252		
FT	HELIX	255	257		
FT	STRAND	266	270		
FT	TURN	275	276		
FT	STRAND	279	282		
FT	HELIX	285	296		
FT	TURN	297	299		
FT	STRAND	302	306		
FT	STRAND	310	319		
FT	TURN	320	321		
FT	STRAND	322	327		
FT	STRAND	345	347		
FT	HELIX	351	360		
FT	TURN	361	362		
FT	HELIX	363	365		
FT	STRAND	369	377		
FT	TURN	378	379		
FT	STRAND	382	388		
FT	TURN	390	391		
FT	TURN	398	398		
FT	HELIX	399	416		
SQ	SEQUENCE	706 AA; 74518 MW; 89373750BF014340 CRC64;			
Query Match 100.0%; Score 45; DB 1; Length 706;					
Best Local Similarity 100.0%; Pred. No. 7.8e+02;					
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1 RPQPPPP 7				
Db	27 RPQPPPP 33				
RESULT 26					
SYN1_MOUSE					
ID	SYN1_MOUSE	STANDARD;	PRT;	706 AA.	
AC	O88935; Q62279; Q8QZT8;				
DT	30-MAY-2000 (Rel. 39, Created)				
DT	13-SEP-2005 (Rel. 48, Last sequence update)				
DT	13-SEP-2005 (Rel. 48, Last annotation update)				
DE	Synapsin-1 (Synapsin I).				
GN	Name=Syn1; Synonyms=Syn-1;				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;				
OC	Muroidea; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM IB).				
RC	STRAIN=C57BL/6; TISSUE=Pancreatic islets;				
RX	MEDLINE=99107854; PubMed=9890964; DOI=10.1074/jbc.274.4.2053;				
RA	Matsumoto K., Ebihara K., Yamamoto H., Tabuchi H., Fukunaga K.,				
RA	Yasunami M., Ohkubo H., Shichiri M., Miyamoto E.;				
RT	"Cloning from insulinoma cells of synapsin I associated with insulin				
RT	secretory granules.";				
RL	J. Biol. Chem. 274:2053-2059(1999).				
RN	[2]				
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 3).				
RC	TISSUE=Eye;				

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 1-125.
RX MEDLINE=94308086; PubMed=8034599;
RA Chin L.S., Li L., Greengard P.;
RT "Neuron-specific expression of the synapsin II gene is directed by a
RT specific core promoter and upstream regulatory elements.";
RL J. Biol. Chem. 269:18507-18513 (1994).
CC -!- FUNCTION: Neuronal phosphoprotein that coats synaptic vesicles,
CC binds to the cytoskeleton, and is believed to function in the
CC regulation of neurotransmitter release. Regulation of
CC neurotransmitter release. The complex formed with NOS1 and CAPON
CC proteins is necessary for specific nitric-oxid functions at a
CC presynaptic level.
CC -!- SUBUNIT: Homodimer. Interacts with CAPON. Forms a ternary complex
CC with NOS1 (By similarity).
CC -!- SUBCELLULAR LOCATION: Synapse.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=Ia;
CC IsoId=O88935-2; Sequence=Displayed;
CC Name=Ib;
CC IsoId=O88935-1; Sequence=VSP_015206, VSP_015207;
CC Name=3;
CC IsoId=O88935-3; Sequence=VSP_015205;
CC Note=No experimental confirmation available;
CC -!- PTM: Substrate of at least four different protein kinases. It is
CC probable that phosphorylation plays a role in the regulation of
CC synapsin-1 in the nerve terminal (By similarity).
CC -!- SIMILARITY: Belongs to the synapsin family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AF085809; AAD09833.1; -; mRNA.
CC EMBL; BC022954; AAH22954.1; -; mRNA.
CC EMBL; L32025; AAA79963.1; -; Genomic_DNA.
CC FIR; A53692; A53692.
CC HSSP; P17599; 1AUX.
CC SMR; O88935; 112-417.
CC Ensembl; ENSMUSG00000037217; Mus musculus.
CC MGI; MGI:98460; Syn1.
CC GO; GO:0045202; C:synapse; IDA.
CC GO; GO:0030672; C:synaptic vesicle membrane; IDA.
CC GO; GO:0046983; F:protein dimerization activity; TAS.
CC InterPro; IPR001359; Synapsin.
CC PANTHER; PTHR10841; Synapsin; 1.
CC Pfam; PF02750; Synapsin_C; 1.
CC Pfam; PF02078; Synapsin_N; 1.
CC PRINTS; PR01368; SYNAPSIN.

DR PROSITE; PS00415; SYNAPSIN_1; 1.
DR PROSITE; PS00416; SYNAPSIN_2; 1.
KW Actin-binding; Alternative splicing; Phosphorylation; Repeat; Synapse.
FT REGION 1 28 A.
FT REGION 29 112 B; linker.
FT REGION 113 420 C; actin-binding and synaptic-vesicle binding.
FT REGION 421 657 D; Pro-rich linker.
FT REGION 658 706 E.
FT MOD_RES 9 Phosphoserine (by CaMK1 and PKA) (By similarity).
FT MOD_RES 568 568 Phosphoserine (by CaMK2) (By similarity).
FT MOD_RES 605 605 Phosphoserine (by CaMK2) (By similarity).
FT VARSPLIC 573 600 Missing (in isoform 3).
FT FTID=VSP_015205.
FT VARSPLIC 662 670 NKSQLTNA -> KASPSQAQP (in isoform Ib).
FT FTID=VSP_015206.
FT VARSPLIC 671 706 Missing (in isoform Ib).
FT FTID=VSP_015207.
FT CONFLICT 44 44 P -> L (in Ref. 3).
SQ SEQUENCE 706 AA; 74097 MW; 04C940E68547372B CRC64;

Query Match 100.0%; Score 45; DB 1; Length 706;
Best Local Similarity 100.0%; Pred. No. 7.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 27 RPQPPPP 33

RESULT 27
Q51L36_MAGGR
ID Q51L36_MAGGR PRELIMINARY; PRT; 721 AA.
AC Q51L36;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=MG03055.4;
OS Magnaporthe grisea 70-15.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthae.
OX NCBI_TaxID=242507;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
RA Bayul T., Blitshsteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA Gnirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamvyselis M., Karlsson E.,
RA Kells C., Kieu A., Kisner P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., Mcghee T., Meldrim J., Meneus L.,
RA Mesirov J., Mihalev A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., News C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
RA O'Neill K., Osman S., Parker S., Perrin D., Phunkhang P., Piquani B.,

RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thoulutsang Y., Topham K.,
RA Towey S., Tsamila T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Magnaporthe grisea.";
RL Submitted (OCT-2003) to the EMBL/GenBank/DBSJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Dean R., Mitchell T., Brown D., Pan H., Thon M.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBSJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Zhu H., Blackmon B.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBSJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBSJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACU01001597; EAA47812.1; -; Genomic_DNA.
DR InterPro; IPR001138; Fungi_Trscrp.N.
DR InterPro; IPR012344; Matrix_HIV/RSV.
DR Pfam; PF00172; Zn_clus; 1.
DR SMART; SM00066; GAL4; 1.
DR PROSITE; PS00463; ZN2_CY6_FUNGAL_1; 1.
DR PROSITE; PS50048; ZN2_CY6_FUNGAL_2; 1.
KW DNA-binding; Hypothetical protein; Metal-binding; Nuclear protein;
KW Transcription; Transcription regulation; Zinc.
SQ SEQUENCE 721 AA; 78599 MW; CODECF788259DDE7 CRC64;

Query Match 100.0%; Score 45; DB 2; Length 721;
Best Local Similarity 100.0%; Pred. No. 7.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 84 RPQPPPP 90

RESULT 28
Q4QF52_LEIMA
ID Q4QF52_LEIMA PRELIMINARY; PRT; 1412 AA.
AC Q4QF52;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Ubiquitin hydrolase, putative (Cysteine peptidase, clan ca, family
DE c19, putative).
GN ORFNames=LmjF15.1300;
OS Leishmania major.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Friedlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Fosker N., Harris D., Oliver K., O'Neil S.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBSJ databases.
DR EMBL; CT005254; CAJ03358.1; -; Genomic_DNA.
KW Hydrolase.
SQ SEQUENCE 1412 AA; 155044 MW; 6657E5CCDD56122D CRC64;

Query Match 100.0%; Score 45; DB 2; Length 1412;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Oy	1 RPQPPPP 7
Db	832 RPQPPPP 838
RESULT 29	
Q96Q04 HUMAN	
ID	Q96Q04_HUMAN PRELIMINARY; PRT; 1480 AA.
AC	Q96Q04;
DT	01-DEC-2001 (TrEMBLrel. 19, Created)
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE	KIAA1883 protein (Fragment).
GN	Name=LMTK3; Synonyms=KIAA1883;
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC	Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	NUCLEOTIDE SEQUENCE.
RC	TISSUE=Brain;
RX	MEDLINE=21456161; PubMed=11572484;
RA	Nagase T., Kikuno R., Ohara O.;
RT	"Prediction of the coding sequences of unidentified human genes. XXI.
RT	The complete sequences of 60 new cDNA clones from brain which code for
RT	large proteins.";
RL	DNA Res. 8:179-187(2001).
DR	EMBL; AB067470; BAB67776.1; -; mRNA.
DR	HSSP; P08069; 1JQH.
DR	Ensembl; ENSG00000142235; Homo sapiens.
DR	HGNC; HGNC:19295; LMTK3.
DR	GO; GO:0005524; F:ATP binding; IEA.
DR	GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR	GO; GO:0016740; F:transferase activity; IEA.
DR	GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR	InterPro; IPR000719; Prot kinase.
DR	InterPro; IPR001245; Tyr_kinase.
DR	InterPro; IPR008266; Tyr_kinase_AS.
DR	PRINTS; PR00109; TYRKINASE.
DR	ProDom; PD000001; Prot kinase; 1.
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR	PROSITE; PS00109; PROTEIN_KINASE_TYR; UNKNOWN_1.
KW	ATP-binding; Kinase; Nucleotide-binding; Transferase;
KW	Tyrosine-protein kinase.
FT	NON_TER 1
SQ	SEQUENCE 1480 AA; 155952 MW; C6137E4A8B484C6F CRC64;
Query Match 100.0%; Score 45; DB 2; Length 1480;	
Best Local Similarity 100.0%; Pred. NO. 1.6e+03;	
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Oy	1 RPQPPPP 7
Db	1162 RPQPPPP 1168
RESULT 30	
MAST2 HUMAN	
ID	MAST2_HUMAN STANDARD; PRT; 1798 AA.
AC	Q6P0Q8; O94899; Q5VT07; Q5VT08; Q7LGC4; Q8NDG1; Q96B94; Q9BYE8;
DT	10-MAY-2005 (Rel. 47, Created)
DT	10-MAY-2005 (Rel. 47, Last sequence update)
DT	13-SEP-2005 (Rel. 48, Last annotation update)
DE	Microtubule-associated serine/threonine-protein kinase 2
DE	(EC 2.7.1.37).
GN	Name=MAST2; Synonyms=KIAA0807, MAST205;
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;

OC	Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), SUBCELLULAR LOCATION, AND
RP	INTERACTION WITH PC-LKC.
RC	TISSUE=Brain;
RX	MEDLINE=22112621; PubMed=12117771; DOI=10.1093/carcin/23.7.1139;
RA	Okazaki N., Takahashi N., Kojima S., Masuho Y., Koga H.;
RT	"Protocadherin LKC, a new candidate for a tumor suppressor of colon
RT	and liver cancers, its association with contact inhibition of cell
RT	proliferation.";
RL	Carcinogenesis 23:1139-1148(2002).
RN	[2]
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA] (ISOFORM 1).
RG	Human chromosome 1 international sequencing consortium;
RL	Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
RN	[3]
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1) AND NUCLEOTIDE
RP	SEQUENCE [LARGE SCALE MRNA] OF 281-1798 (ISOFORM 2).
RC	TISSUE=Lymph, and Placenta;
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA	Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA	Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA	Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT	"Generation and initial analysis of more than 15,000 full-length human
RT	and mouse cDNA sequences.";
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN	[4]
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 470-1797 (ISOFORM 1).
RC	TISSUE=Brain;
RX	MEDLINE=99087487; PubMed=9872452;
RA	Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Miyajima N.,
RA	Tanaka A., Kotani H., Nomura N., Ohara O.;
RT	"Prediction of the coding sequences of unidentified human genes. XI.
RT	The complete sequences of 100 new cDNA clones from brain which code
RT	for large proteins in vitro.";
RL	DNA Res. 5:277-286(1998).
RN	[5]
RP	SEQUENCE REVISION TO C-TERMINUS.
RA	Ohara O., Suyama M., Nagase T., Ishikawa K.-I., Kikuno R.;
RL	Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
RN	[6]
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 833-1797 (ISOFORM 1).
RC	TISSUE=Testis;
RG	The German cDNA consortium;
RL	Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
RN	[7]
RP	TISSUE SPECIFICITY.
RX	PubMed=8902215;
RA	Walden P.D., Millette C.F.;
RT	"Increased activity associated with the MAST205 protein kinase complex
RT	during mammalian spermiogenesis.";
RL	Biol. Reprod. 55:1039-1044(1996).
CC	-1- FUNCTION: Appears to link the dystrophin/utrophin network with
CC	microtubule filaments via the syntrophins. Phosphorylation of DMD
CC	or UTRN may modulate their affinities for associated proteins.
CC	Functions in a multi-protein complex in spermatid maturation.
CC	Regulates lipopolysaccharide-induced IL-12 synthesis in
CC	macrophages by forming a complex with TRAF6, resulting in the

CC inhibition of TRAF6 NF-kappaB activation (By similarity).

CC -|- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.

CC -|- COPACTOR: Magnesium (By similarity).

CC -|- INTERACTION:

CC Q9BYE9:PC-LKC; NbExp=3; IntAct=EBI-493777, EBI-493793;

CC -|- SUBCELLULAR LOCATION: Cytoplasmic. Recruited to the sub-membranous area on interaction with PC-LKC.

CC -|- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=2;

CC Name=1;

CC IsoId=Q6P0Q8-1; Sequence=Displayed;

CC Name=2;

CC IsoId=Q6P0Q8-2; Sequence=VSP_051698, VSP_051699, VSP_051700;

CC Note=No experimental confirmation available;

CC -|- TISSUE SPECIFICITY: Abundant in the testis.

CC -|- PTM: Phosphorylated and ubiquitinated. N-terminal ubiquitination leads to degradation of MAST2 by proteasome-mediated proteolysis. N-terminal phosphorylation appears to be a prerequisite for ubiquitination (By similarity).

CC -|- SIMILARITY: Belongs to the Ser/Thr protein kinase family.

CC -|- SIMILARITY: Contains 1 PDZ (DHR) domain.

CC -|- CAUTION: Ref.1 sequence differs from that shown due to a frameshift in position 1688.

CC -|- CAUTION: Ref.2 (CAH73245, CAI16563 and CAI21706) sequences differ from that shown due to erroneous gene model prediction.

CC -|- CAUTION: Ref.3 (AAH15816) sequence differs from that shown due to frameshifts in positions 1114 and 1387.

CC -----

CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

CC -----

DR EMBL; AB047005; BAB40778.1; ALT_FRAME; mRNA.

DR EMBL; AL645480; CAH73244.1; -; Genomic DNA.

DR EMBL; AL358075; CAH73244.1; JOINED; Genomic DNA.

DR EMBL; AL603882; CAH73244.1; JOINED; Genomic DNA.

DR EMBL; AL603888; CAH73244.1; JOINED; Genomic DNA.

DR EMBL; AL645480; CAH73245.1; ALT_SEQ; Genomic DNA.

DR EMBL; AL358075; CAH73245.1; JOINED; Genomic DNA.

DR EMBL; AL603888; CAH73245.1; JOINED; Genomic DNA.

DR EMBL; AL603882; CAI16217.1; -; Genomic DNA.

DR EMBL; AL358075; CAI16217.1; JOINED; Genomic DNA.

DR EMBL; AL603888; CAI16217.1; JOINED; Genomic DNA.

DR EMBL; AL645480; CAI16217.1; JOINED; Genomic DNA.

DR EMBL; AL603888; CAI16562.1; -; Genomic DNA.

DR EMBL; AL358075; CAI16562.1; JOINED; Genomic DNA.

DR EMBL; AL603882; CAI16562.1; JOINED; Genomic DNA.

DR EMBL; AL645480; CAI16562.1; JOINED; Genomic DNA.

DR EMBL; AL603888; CAI16563.1; ALT_SEQ; Genomic DNA.

DR EMBL; AL358075; CAI16563.1; JOINED; Genomic DNA.

DR EMBL; AL645480; CAI16563.1; JOINED; Genomic DNA.

DR EMBL; AL358075; CAI21705.1; -; Genomic DNA.

DR EMBL; AL603882; CAI21705.1; JOINED; Genomic DNA.

DR EMBL; AL603888; CAI21705.1; JOINED; Genomic DNA.

DR EMBL; AL645480; CAI21705.1; JOINED; Genomic DNA.

DR EMBL; AL358075; CAI21706.1; ALT_SEQ; Genomic DNA.

DR EMBL; AL603888; CAI21706.1; JOINED; Genomic DNA.

DR EMBL; AL645480; CAI21706.1; JOINED; Genomic DNA.

DR EMBL; BC015816; AAH15816.2; ALT_INIT; mRNA.

DR EMBL; BC065499; AAH65499.1; -; mRNA.

DR EMBL; AB018350; BAA34527.2; -; mRNA.

DR EMBL; AL833919; CAD38775.1; -; mRNA.

DR HSSP; P05132; 1APM.

DR IntAct; Q6P0Q8; -.

DR Ensembl; ENSG00000086015; Homo sapiens.

DR HGNC; HGNC:19035; MAST2.

DR GO; GO:0005524; F:ATP binding; ISS.

DR GO; GO:0000287; F:magnesium ion binding; ISS.

DR GO; GO:0006468; P:protein amino acid phosphorylation; ISS.

DR GO; GO:0045075; P:regulation of interleukin-12 biosynthesis; ISS.

DR GO; GO:0048515; P:spermatid differentiation; ISS.

DR InterPro; IPR001478; PDZ.

DR InterPro; IPR000961; Pkinase C.

DR InterPro; IPR000719; Prot_kinase.

DR InterPro; IPR008271; Ser_thr_pkin_AS.

DR InterPro; IPR002290; Ser_thr_pkinase.

DR Pfam; PF00595; PDZ; 1.

DR Pfam; PF00069; Pkinase; 1.

DR Pfam; PF00433; Pkinase C; 1.

DR ProDom; PD000001; Prot_kinase; 1.

DR SMART; SM00228; PDZ; 1.

DR SMART; SM00220; S_TKC; 1.

DR PROSITE; PS50106; PDZ; 1.

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.

DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.

KW Alternative splicing; ATP-binding; Cytoskeleton; Kinase; Magnesium; Membrane; Metal-binding; Nucleotide-binding; Phosphorylation; Serine/threonine-protein kinase; Transferase; Ubl conjugation.

KW Serine/threonine-protein kinase; Transferase; Ubl conjugation.

FT DOMAIN 512 785 Protein kinase.

FT DOMAIN 1104 1192 PDZ.

FT NP_BIND 518 526 ATP (By similarity).

FT ACT_SITE 635 635 Proton acceptor (By similarity).

FT BINDING 541 541 ATP (By similarity).

FT VARSPLIC 327 396 Missing (in isoform 2).

FT VARSPLIC 1091 1113 Missing (in isoform 2).

FT VARSPLIC 1290 1386 Missing (in isoform 2).

FT CONFLICT 388 388 D -> E (in Ref. 1).

FT CONFLICT 659 659 I -> M (in Ref. 1, 3; AAH15816 and 4).

FT CONFLICT 1225 1225 Missing (in Ref. 3; AAH65499).

FT CONFLICT 1551 1551 D -> G (in Ref. 1 and 4).

SQ SEQUENCE 1798 AA; 196436 MW; DF5C92078A3451AF CRC64;

Query Match 100.0%; Score 45; DB 1; Length 1798;

Best Local Similarity 100.0%; Pred. No. 1.9e+03;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7

Db 9 RPQPPPP 15

RESULT 31

Q9ATK5_CHLRE PRELIMINARY; PRT; 2301 AA.

AC Q9ATK5;

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE PF6 protein.

OS Chlamydomonas reinhardtii.

OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae;

OC Chlamydomonadales; Chlamydomonadaceae; Chlamydomonas.

OX NCBI_TaxID=3055;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=21gr;

RX MEDLINE=21149846; PubMed=11251084;

RA Rupp G., O'Toole E., Porter M.E.;

RT "The Chlamydomonas PF6 locus encodes a large alanine/proline-rich polypeptide that is required for assembly of a central pair projection and regulates flagellar motility.";

RL Mol. Biol. Cell 12:739-751(2001).

DR EMBL; AF327876; AAK38270.1; -; Genomic DNA.

SQ SEQUENCE 2301 AA; 237480 MW; A296943A4CFE9F9E CRC64;

Query Match 100.0%; Score 45; DB 2; Length 2301;

Best Local Similarity 100.0%; Pred. No. 2.5e+03;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7

```
Db          1266 RPQPPP 1272
|||||
RESULT 32
MLL2_HUMAN
ID   MLL2_HUMAN          STANDARD;          PRT;   5262 AA.
AC   O14686; O14687;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Myeloid/lymphoid or mixed-lineage leukemia protein 2 (ALL1-related
DE   protein).
GN   Name=MLL2; Synonyms=ALR;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE (ISOFORMS 1; 2 AND 3).
RX   MEDLINE=97388474; PubMed=9247308; DOI=10.1038/sj.onc.1201211;
RA   Prasad R., Zhadanov A.B., Sedkov Y., Bullrich F., Druck T.,
RA   Rallapalli R., Yano T., Alder H., Croce C.M., Huebner K., Mazo A.,
RA   Canaani E.;
RT   "Structure and expression pattern of human ALR, a novel gene with
RT   strong homology to ALL-1 involved in acute leukemia and to Drosophila
RT   trithorax."
RL   Oncogene 15:549-560(1997).
RN   [2]
RP   INTERACTION WITH ASC-2/NCOA6 CONTAINING COMPLEX.
RC   TISSUE=Cervical carcinoma;
RX   MEDLINE=22371496; PubMed=12482968; DOI=10.1128/MCB.23.1.140-149.2003;
RA   Goo Y.-H., Sohn Y.C., Kim D.-H., Kim S.-W., Kang M.-J., Jung D.-J.,
RA   Kwak E., Barlev N.A., Berger S.L., Chow V.T., Roeder R.G.,
RA   Azorsa D.O., Meltzer P.S., Suh P.-G., Song E.J., Lee K.-J., Lee Y.C.,
RA   Lee J.W.;
RT   "Activating signal cointegrator 2 belongs to a novel steady-state
RT   complex that contains a subset of trithorax group proteins."
RL   Mol. Cell. Biol. 23:140-149(2003).
RN   [3]
RP   PHOSPHORYLATION SITE SER-4463.
RX   PubMed=15302935; DOI=10.1073/pnas.0404720101;
RA   Beausoleil S.A., Jedrychowski M., Schwartz D., Elias J.E., Villen J.,
RA   Li J., Cohn M.A., Cantley L.C., Gygi S.P.;
RT   "Large-scale characterization of HeLa cell nuclear phosphoproteins."
RL   Proc. Natl. Acad. Sci. U.S.A. 101:12130-12135(2004).
CC   -1- FUNCTION: May be involved in transcriptional regulation.
CC   -1- SUBUNIT: Belongs to the ASC-2/NCOA6 complex (ASCOM), which
CC   contains ASC-2/NCOA6, the retinoblastoma-binding protein RBQ-3/
CC   RBBP5, alpha- and beta-tubulins, the trithorax group proteins MLL2
CC   and MLL3, and ASH2/ASCL2.
CC   -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC   -1- ALTERNATIVE PRODUCTS:
CC   Event=Alternative splicing; Named isoforms=3;
CC   Name=1;
CC   IsoId=O14686-1; Sequence=Displayed;
CC   Name=2;
CC   IsoId=O14686-2; Sequence=VSP_008563, VSP_008559;
CC   Name=3;
CC   IsoId=O14686-3; Sequence=VSP_008560;
CC   TISSUE SPECIFICITY: Expressed in most adult tissues, including a
CC   variety of hematopoietic cells, with the exception of the liver.
CC   -1- MISCELLANEOUS: This gene mapped to a chromosomal region involved
CC   in duplications and translocations associated with cancer.
CC   -1- SIMILARITY: Belongs to the transcription factor trithorax family.
CC   -1- SIMILARITY: Contains 5 PHD-type zinc fingers.
CC   -1- SIMILARITY: Contains 1 post-SET domain.
CC   -1- SIMILARITY: Contains 1 RING-type zinc finger.
CC   -1- SIMILARITY: Contains 1 SET domain.
-----
This Swiss-prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
-----
```

```
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AF010403; AAC51734.1; -; mRNA.
DR EMBL; AF010404; AAC51735.1; -; mRNA.
DR PIR; T03454; T03454.
DR PIR; T03455; T03455.
DR HSSP; Q14839; 1MM2.
DR Ensembl; ENSG00000167548; Homo sapiens.
DR HGNC; HGNC:7133; MLL2.
DR MIM; 602113; -.
DR GO; GO:0005634; C:nucleus; NAS.
DR GO; GO:0003700; F:transcription factor activity; TAS.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
DR GO; GO:0006366; P:transcription from RNA polymerase II promoter; TAS.
DR InterPro; IPR003889; FYrich_C.
DR InterPro; IPR003888; FYrich_N.
DR InterPro; IPR003616; PostSET.
DR InterPro; IPR001214; SET.
DR InterPro; IPR001965; Znf_PHD.
DR InterPro; IPR001841; Znf_RING.
DR Pfam; PF05965; FYRC; 1.
DR Pfam; PF05964; FYRN; 1.
DR Pfam; PF00628; PHD; 5.
DR Pfam; PF00856; SET; 1.
DR PROSITE; PS00868; POST_SET; 1.
DR PROSITE; PS0280; SET; 1.
DR PROSITE; PS01359; ZF_PHD_1; 5.
DR PROSITE; PS0016; ZF_PHD_2; 5.
DR PROSITE; PS0089; ZF_RING_2; 1.
KW Alternative splicing; Coiled coil; Metal-binding; Nuclear protein;
KW Phosphorylation; Polymorphism; Repeat; Transcription;
KW Transcription regulation; Zinc; Zinc-finger.
FT REPEAT 442 446 1.
FT REPEAT 460 464 2.
FT REPEAT 469 473 3.
FT REPEAT 496 500 4.
FT REPEAT 504 508 5.
FT REPEAT 521 525 6.
FT REPEAT 555 559 7.
FT REPEAT 564 568 8.
FT REPEAT 573 577 9.
FT REPEAT 582 586 10.
FT REPEAT 609 613 11.
FT REPEAT 618 622 12.
FT REPEAT 627 631 13.
FT REPEAT 645 649 14.
FT REPEAT 663 667 15.
FT DOMAIN 5121 5242 SET.
FT DOMAIN 5246 5262 Post-SET.
FT ZN_FING 226 276 PHD-type 1.
FT ZN_FING 229 274 RING-type.
FT ZN_FING 273 323 PHD-type 2.
FT ZN_FING 1102 1155 PHD-type 3.
FT ZN_FING 1152 1202 PHD-type 4.
FT ZN_FING 1229 1284 PHD-type 5.
FT REGION 439 668 15 X 5 AA repeats of S/P-P-P-E/P-E/A.
FT COILED 2397 2436 Potential.
FT COILED 2788 2809 Potential.
FT COILED 2974 3001 Potential.
FT COILED 3286 3342 Potential.
FT COILED 3437 3476 Potential.
FT COILED 3621 3701 Potential.
FT COILED 4265 4287 Potential.
FT COMPBIAS 229 326 Cys-rich.
FT COMPBIAS 374 922 Pro-rich.
FT COMPBIAS 1015 1053 Arg-rich.
FT COMPBIAS 1122 1235 Cys-rich.
FT COMPBIAS 1832 2351 Pro-rich.
FT COMPBIAS 2536 2547 Gln-rich.
FT COMPBIAS 2587 2703 Pro-rich.
FT COMPBIAS 2986 4000 Gln-rich.
```

FT	COMPBIAS	3966	4085	Pro-rich.
FT	COMPBIAS	4634	4702	Pro-rich.
FT	MOD_RES	4463	4463	Phosphoserine.
FT	VARSPLIC	1	305	Missing (in isoform 2).
FT				/FTid=VSP 008563.
FT	VARSPLIC	306	672	PMEELPAHSWKACRVCRACGAGSAELNPNSEWFENYSLC
FT				HRCHKAQGGQTIKRSVAEQHTPVCRSFSPPEPGDTPTDEPDA
FT				LYVACQGPQKGGHVTSMQPKFPGPLQCEAKPLGKAGVQLEP
FT				QLEAPLNEEMPLLLPPEESLPQPEESLPPEESPTSPPEASRLSP
FT				PPEELPASPLPEALHLSRPLEESPLSPPEESPLSPPEESS
FT				PFSPLEESPLSPPEESPPSPALETPLSPPEASPLSPPEEF
FT				SPLSPPEELTSPPEASRLSPPEESPLSPPEASRLSPPEESMSP
FT				PPEASRLFPPEESPLSPPEESPLSPPEASRLSPPEESPLSP
FT				PMSPPEESPMSPPEEVSRSLSPVVSRLSPPEESPLSP
FT				-> MSPPPEESPMSPPEASRLFPPEESPLSPPEESPLSP
FT				PPPEASRLSPPEESPMSPPEESPMSPPEEVSRSLSPVVS
FT				SRLSPPEESPLSPPEESPTSPPEASRLSPPEESPTSP
FT				PPEDSPASPPEDSLMSLPLEESPLLPPEEPQLCPRSEGP
FT				HLSPRPEEHLSPRPEEHLSPQAEPEHLSPQPEEPCLCVAV
FT				PEEPHLSPQAEGPLHLSQPPEELHLSQTEEPHLSPVPEEPC
FT				LSPQPEESHLSPQSEEPCLSPRPEESHLSPLEKPPLSRP
FT				EKPPEEGQCPAPEELPLFPPPGEPRLSPLLGEPALSEPGE
FT				PPLSPLPEELPLSPSGEPLSPQLMPPDPLPPPLSPIITAA
FT				A (in isoform 2).
FT				/FTid=VSP 008559.
FT	VARSPLIC	1454	1454	E -> EGEG_ (in isoform 3).
FT				/FTid=VSP 008560.
FT	VARIANT	4949	4949	R -> H (in dbSNP:3782356).
FT				/FTid=VAR 017115.
FT	SEQUENCE	5262 AA;	564185 MW;	26B7C74CAD417E44 CRC64;
Query Match 100.0%; Score 45; DB 1; Length 5262;				
Best Local Similarity 100.0%; Pred. No. 5.6e+03;				
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1	RQPPPPP 7		
Db	2111	RQPPPPP 2117		
RESULT 33				
Q67YU2	ARATH			
ID	Q67YU2_ARATH	PRELIMINARY;	PRT;	107 AA.
AC	Q67YU2;			
DT	25-OCT-2004	(TrEMBLrel. 28, Created)		
DT	25-OCT-2004	(TrEMBLrel. 28, Last sequence update)		
DT	25-OCT-2004	(TrEMBLrel. 28, Last annotation update)		
DE	Hypothetical protein At4g21720.			
GN	Name=At4g21720;			
OS	Arabidopsis thaliana (Mouse-ear cross).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;			
OC	eurosids II; Brassicales; Brassicaceae; Arabidopsi.			
OX	NCBI_TaxID=3702;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RA	Totoki Y., Seki M., Ishida J., Nakajima M., Enju A., Kamiya A.,			
RA	Narusaka M., Shin-i T., Nakagawa M., Sakamoto N., Oishi K., Kohara Y.,			
RA	Kobayashi M., Toyoda A., Sakaki Y., Sakurai T., Iida K., Akiyama K.,			
RA	Satou M., Toyoda T., Konagaya A., Carninci P., Kawai J.,			
RA	Hayashizaki Y., Shinozaki K.;			
RT	"Large-scale analysis of RIKEN Arabidopsis full-length (RAFL) cDNAs.";			
RL	Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.			
DR	EMBL;	AK176376;	BAD44139.1;	-; mRNA.
KW	Hypothetical protein.			
SQ	SEQUENCE	107 AA;	11687 MW;	9A48D4B53F4159F4 CRC64;
Query Match 93.3%; Score 42; DB 2; Length 107;				
Best Local Similarity 85.7%; Pred. No. 2.8e+02;				
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;				
QY	1	RQPPPPP 7		

Db	81	KPQPPPP 87		
RESULT 34				
Q9SVS5	ARATH			
ID	Q9SVS5_ARATH	PRELIMINARY;	PRT;	139 AA.
AC	Q9SVS5;			
DT	01-MAY-2000	(TrEMBLrel. 13, Created)		
DT	01-MAY-2000	(TrEMBLrel. 13, Last sequence update)		
DT	01-FEB-2005	(TrEMBLrel. 29, Last annotation update)		
DE	Hypothetical protein F17L22.180 (Hypothetical protein AT4g21720).			
GN	Name=F17L22.180; Synonyms=AT4g21720, At4g21720;			
OS	Arabidopsis thaliana (Mouse-ear cross).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;			
OC	eurosids II; Brassicales; Brassicaceae; Arabidopsi.			
OX	NCBI_TaxID=3702;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RA	Bevan M., Van Der Schueren J., Chuang Y-J., Voet M., Robben J.,			
RA	Volckaert G., Bancroft I., Mewes H.W., Mayer K.F.X., Schueller C.;			
RL	Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	NUCLEOTIDE SEQUENCE.			
RA	Van Der Schueren J., Vandenbussche F., Chuang Y.J., Braeken M.,			
RA	Robben J., Volckaert G., Mewes H.W., Lemcke K., Mayer K.F.X.;			
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	NUCLEOTIDE SEQUENCE.			
RA	EU Arabidopsis sequencing project;			
RL	Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	NUCLEOTIDE SEQUENCE.			
RA	Cheuk R., Chen H., Kim C.J., Shinn P., Carninci P., Hayashizaki Y.,			
RA	Ishida J., Kamiya A., Kawai J., Narusaka M., Sakurai T., Satou M.,			
RA	Seki M., Shinozaki K., Ecker J.R.;			
RL	Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.			
RN	[6]			
RP	NUCLEOTIDE SEQUENCE.			
RA	Totoki Y., Seki M., Ishida J., Nakajima M., Enju A., Kamiya A.,			
RA	Narusaka M., Shin-i T., Nakagawa M., Sakamoto N., Oishi K., Kohara Y.,			
RA	Kobayashi M., Toyoda A., Sakaki Y., Sakurai T., Iida K., Akiyama K.,			
RA	Satou M., Toyoda T., Konagaya A., Carninci P., Kawai J.,			
RA	Hayashizaki Y., Shinozaki K.;			
RT	"Large-scale analysis of RIKEN Arabidopsis full-length (RAFL) cDNAs.";			
RL	Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.			
DR	EMBL;	AL035527;	CAB36816.1;	-; Genomic_DNA.
DR	EMBL;	AL161555;	CAB81279.1;	-; Genomic_DNA.
DR	EMBL;	BT010603;	AAQ89625.1;	-; mRNA.
DR	EMBL;	AK175223;	BAD42986.1;	-; mRNA.
DR	PIR;	T05847;		
KW	Hypothetical protein.			
SQ	SEQUENCE	139 AA;	15436 MW;	9C01143EF81528A8 CRC64;
Query Match 93.3%; Score 42; DB 2; Length 139;				
Best Local Similarity 85.7%; Pred. No. 3.6e+02;				
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;				
QY	1	RQPPPPP 7		
Db	113	KPQPPPP 119		
RESULT 35				
Q5N912	ORYSA			
ID	Q5N912_ORYSA	PRELIMINARY;	PRT;	169 AA.
AC	Q5N912;			
DT	01-FEB-2005	(TrEMBLrel. 29, Created)		
DT	01-FEB-2005	(TrEMBLrel. 29, Last sequence update)		

DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein P0683F02.11.
GN Name=P0683F02.11;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Niumura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
RA Ikeno M., Itoh S., Itoh T., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,
RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA Yano M., Jiang J., Gojobori T.;
RT "The genome sequence and structure of rice chromosome 1.";
RL Nature 420:312-316(2002).
DR EMBL; AP003289; BAD82047.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 169 AA; 17621 MW; AC7AE90E8E6AE196 CRC64;

Query Match 93.3%; Score 42; DB 2; Length 169;
Best Local Similarity 85.7%; Pred. No. 4.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 95 KPQPPPP 101

RESULT 36
Q6Z0S6 ORYSA
ID Q6Z0S6_ORYSA PRELIMINARY; PRT; 246 AA.
AC Q6Z0S6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Far-red impaired response-like protein.
GN Name=P0439B07.7; Synonyms=OJ1125_C01.31;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 8, PAC
RT clone:P0439B07.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 8, BAC
RT clone:OJ1125_C01.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJJ databases.
DR EMBL; AP005480; BAD10385.1; -; Genomic_DNA.
DR EMBL; AP004044; BAD33021.1; -; Genomic_DNA.
DR Gramene; Q6Z0S6; -.
SQ SEQUENCE 246 AA; 26493 MW; 7057C07BBBFFFD80 CRC64;

Query Match 93.3%; Score 42; DB 2; Length 246;
Best Local Similarity 85.7%; Pred. No. 6.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 203 RPEPPPP 209

RESULT 37
Q6ZQX5 HUMAN
ID Q6ZQX5_HUMAN PRELIMINARY; PRT; 279 AA.
AC Q6ZQX5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein FLJ46823.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Uterus;
RA Suzuki O., Sasaki N., Aotsuka S., Shoji T., Ichihara T., Shiohata N.,
RA Matsumoto K., Hirano M., Sano S., Nomura R., Yoshikawa Y.,
RA Matsumura Y., Moriya S., Chiba E., Momiyama H., Onogawa S.,
RA Kaeriyama S., Satoh N., Matsunawa H., Takahashi E., Kataoka R.,
RA Kuga N., Kuroda A., Satoh I., Kamata K., Takami S., Terashima Y.,
RA Watanabe M., Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A.,
RA Ishii S., Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Yamashita H., Matsuo K., Nakamura Y., Sekine M.,
RA Kikuchi H., Kanda K., Wagatsuma M., Murakawa K., Kanehori K.,
RA Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y.,
RA Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJJ databases.
DR EMBL; AK128663; BAC87557.1; -; mRNA.
SQ SEQUENCE 279 AA; 30713 MW; E6200AEAOA309820 CRC64;

Query Match 93.3%; Score 42; DB 2; Length 279;
Best Local Similarity 85.7%; Pred. No. 7.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 21 RPEPPPP 27

RESULT 38
Q96KT3 HUMAN
ID Q96KT3_HUMAN PRELIMINARY; PRT; 287 AA.
AC Q96KT3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein C8orf7 (Fragment).
GN Name=C8orf7;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Appel S., Bergheim A., Reichwald K., Reis A., Rosenthal A., Ramsay M.,
RA Hennies H.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJJ databases.
DR EMBL; AJ301560; CAC82736.1; -; mRNA.
DR Ensembl; ENSG00000164724; Homo sapiens.
DR HGNC; HGNC:15543; C8orf7.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 287 AA; 30400 MW; 39234688A8B01A69 CRC64;

Query Match 93.3%; Score 42; DB 2; Length 287;

Best Local Similarity 85.7%; Pred. No. 7.4e+02; Mismatches 0; Indels 0; Gaps 0; Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
||:||||
Db 114 RPEPPPP 120

RESULT 39
Q7XPK1_ORYSA
ID Q7XPK1_ORYSA PRELIMINARY; PRT; 309 AA.
AC Q7XPK1;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE OSJNBa0087O24.9 protein.
GN Name=OSJNBa0087O24.9;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartioideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]

RP NUCLEOTIDE SEQUENCE.
RX PubMed=12447439; DOI=10.1038/nature01183;
RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,
RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,
RA Weng Q., Zhang L., Lu Y., Mu J., Lu Y., Zhang L.S., Yu Z., Fan D.,
RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,
RA Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,
RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,
RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,
RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
RA Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,
RA Han B.;
RT "Sequence and analysis of rice chromosome 4.";
RL Nature 420:316-320(2002).
DR EMBL; AL606646; CAB03586.1; -; Genomic_DNA.
DR Gramene; Q7XPK1; -.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR004087; KH.
DR InterPro; IPR004088; KH type 1.
DR InterPro; IPR000571; Znf_CCCH.
DR Pfam; PF00013; KH 1; 1.
DR Pfam; PF00642; zf-CCCH; 3.
DR SMART; SM00322; KH; 1.
DR SMART; SM00356; Znf_C3H1; 3.
DR PROSITE; PS50084; KH TYPE 1; 1.
SQ SEQUENCE 309 AA; 31788 MW; D76FF408DB12DD81 CRC64;

Query Match 93.3%; Score 42; DB 2; Length 309;
Best Local Similarity 85.7%; Pred. No. 8e+02; Mismatches 0; Indels 0; Gaps 0; Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
||:||||
Db 156 RPEPPPP 162

RESULT 40
Q8UDG0_AGRT5
ID Q8UDG0_AGRT5 PRELIMINARY; PRT; 320 AA.
AC Q8UDG0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein Atu2166.
GN OrderedLocusNames=Atu2166;
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Dupont;
RX MEDLINE=21608550; PubMed=11743193; DOI=10.1126/science.1066804;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A., Gordon D.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Perry M.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Dolan M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens C58.";
RL Science 294:2317-2323(2001).
DR EMBL; AE009166; AAL43155.1; -; Genomic_DNA.
DR PIR; AE2842; AE2842.
DR PIR; F97619; F97619.
DR InterPro; IPR009273; DUF930.
DR Pfam; PF06059; DUF930; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 320 AA; 34447 MW; B39449CD3D1FDBD6 CRC64;

Query Match 93.3%; Score 42; DB 2; Length 320;
Best Local Similarity 85.7%; Pred. No. 8.3e+02; Mismatches 0; Indels 0; Gaps 0; Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
:|||||
Db 74 KPQPPPP 80

RESULT 41
Q4IEH9_GIBZE
ID Q4IEH9_GIBZE PRELIMINARY; PRT; 355 AA.
AC Q4IEH9;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=FG04379.1;
OS Gibberella zeae PH-1.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
OX NCBI_TaxID=229533;
RN [1]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PH-1;
RA Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavkiy L.,
RA Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA Choepe J.Y., Collymore A., Cook A., Cooke P., Corum B., DeArellano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neill D.,
RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA Roman J., Schauer S., Schupback R., Seaman S., Severy P., Smirnov S.,
RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA Lander E.;
RT "Fusarium graminearum genome sequence.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBSJ databases.

```
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACM01000190; EAA72760.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 355 AA; 39535 MW; 3D967546B01D90A5 CRC64;

Query Match 93.3%; Score 42; DB 2; Length 355;
Best Local Similarity 85.7%; Pred. No. 9.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 186 KPQPPPP 192

RESULT 42
Q7CXQ3 AGRT5
ID Q7CXQ3_AGR75 PRELIMINARY; PRT; 359 AA.
AC Q7CXQ3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE AGR_C_3928p.
GN OrderedLocusNames=AGR_C_3928;
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Cereon;
RX MEDLINE=21608551; PubMed=11743194; DOI=10.1126/science.1066803;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houniel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
DR EMBL; AE008132; AAK87911.1; -; Genomic_DNA.
DR InterPro; IPR009273; DUF930.
DR Pfam; PF06059; DUF930; 1.
DR SEQUENCE 359 AA; 39061 MW; C6CFFD657C439846 CRC64;

Query Match 93.3%; Score 42; DB 2; Length 359;
Best Local Similarity 85.7%; Pred. No. 9.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 113 KPQPPPP 119

RESULT 43
Q4SJ81 TETNG
ID Q4SJ81_TETNG PRELIMINARY; PRT; 377 AA.
AC Q4SJ81;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 4 SCAF14575, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00017324001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontoidea; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
```

```
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Guzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff JN., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAE01014575; CAF99301.1; -; Genomic_DNA.
FT NON_TER 1 1
FT NON_TER 377 377
SQ SEQUENCE 377 AA; 43608 MW; A272994E5E8C5770 CRC64;

Query Match 93.3%; Score 42; DB 2; Length 377;
Best Local Similarity 85.7%; Pred. No. 9.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 82 RPEPPPP 88

RESULT 44
Q61C10 CAEBR
ID Q61C10_CAEBR PRELIMINARY; PRT; 464 AA.
AC Q61C10;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG13131 (Fragment).
GN Name=CBG13131;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG The C.briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAC01000062; CAB67589.1; -; Genomic_DNA.
DR InterPro; IPR010480; Prot_inh_I33.
DR Pfam; PF06394; Pepsin-I3; 1.
KW Hypothetical protein.
FT NON_TER 464 464
SQ SEQUENCE 464 AA; 52612 MW; 9B6F52C4CEF5E0F3 CRC64;

Query Match 93.3%; Score 42; DB 2; Length 464;
Best Local Similarity 85.7%; Pred. No. 1.2e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 235 KPQPPPP 241
```



```

RESULT 45
Q9UVD1_PNECA
ID Q9UVD1_PNECA PRELIMINARY; PRT; 493 AA.
AC Q9UVD1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Kexin-like serine endoprotease (Fragment).
OS Pneumocystis carinii.
OC Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes; Pneumocystidaceae;
OC Pneumocystis.
OX NCBI_TaxID=4754;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=rat;
RA Russian D.A., Andrawia-Sorral V., Kovacs J.A.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF009222; AAF21601.1; -; mRNA.
DR HSSP; P13134; IOT5.
DR MEROPS; S08.011; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0005199; F:structural constituent of cell wall; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002029; Pept_S8_S53.
DR InterPro; IPR003882; Piatil_extensin.
DR InterPro; IPR002884; Pprtnconvertsp.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF01483; P_proprotein; 1.
DR PRINTS; PR01218; PSTEXTENSIN.
DR ProDom; PD000717; Pprtnconvertsp; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Protease.
FT NON TER 1
SQ SEQUENCE 493 AA; 54145 MW; 2D896F9D94155C0D CRC64;

Query Match 93.3%; Score 42; DB 2; Length 493;
Best Local Similarity 85.7%; Pred. No. 1.3e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 401 KPQPPPP 407

RESULT 46
Q9RJX6_STRCO
ID Q9RJX6_STRCO PRELIMINARY; PRT; 496 AA.
AC Q9RJX6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative export protein.
GN OrderedLocusNames=SC01194; ORFNames=SCG11A.25;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neil S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
"Complete genome sequence of the model actinomycete Streptomyces

```

```

RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939108; CAB61606.1; -; Genomic_DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0015520; F:tetracycline:hydrogen antiporter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0046677; P:response to antibiotic; IEA.
DR GO; GO:0015904; P:tetracycline transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR004638; Efflux_EmrB.
DR InterPro; IPR001220; Lectin_legB.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR011701; MFS_1.
DR InterPro; IPR001411; TCR_TetB.
DR Pfam; PF07690; MFS_1; 1.
DR PRINTS; PR01036; TCR_TETB.
DR TIGRPAMS; TIGR00711; efflux_EmrB; 1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
DR PROSITE; PS50850; MFS; 1.
KW Complete proteome.
SQ SEQUENCE 496 AA; 50938 MW; 61CD995728B8589E CRC64;

Query Match 93.3%; Score 42; DB 2; Length 496;
Best Local Similarity 85.7%; Pred. No. 1.3e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 473 RPEPPPP 479

RESULT 47
Q6I587_ORYSA
ID Q6I587_ORYSA PRELIMINARY; PRT; 542 AA.
AC Q6I587;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative calcium-dependent protein kinase.
GN Name=OSJNBa0009C07.9;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,
RA Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,
RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C., Leu H.-L.,
RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
RA Wu H.-P., Shaw J.-F.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: May play a role in signal transduction pathways that
CC involve calcium as a second messenger (By similarity).
CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AC137608; AAT47064.1; -; Genomic_DNA.
DR Gramene; Q6I587; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011992; EF-Hand_type.
DR InterPro; IPR002048; EF_hand_Ca_bd.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00036; ehand; 4.
DR ProDom; PD000012; EF-hand; 2.

```

```
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00054; EFh; 4.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00018; EF_HAND; 4.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Calcium; Kinase; Nucleotide-binding; Repeat;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 542 AA; 61374 MW; 5E079B20A858B519 CRC64;

Query Match          93.3%; Score 42; DB 2; Length 542;
Best Local Similarity 85.7%; Pred. No. 1.4e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 41 KPQPPPP 47

RESULT 48
Q9FMF8_ARATH PRELIMINARY; PRT; 549 AA.
AC Q9FMF8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE 2-oxoglutarate/malate translocator.
GN Name=At5g64280;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98162728; PubMed=9501997;
RA Nakamura Y., Sato S., Kaneko T., Kotani H., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. III.
RT Sequence features of the regions of 1,191,918 bp covered by seventeen
RT physically assigned P1 clones.";
RL DNA Res. 4:401-414(1997).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hsuan V.W.,
RA Lee J.M., Quach H.L., Tang C., Toriumi M., Wallender E.K., Wong C.,
RA Wu H.C., Yu G., Yuan S., Bowser L., Carninci P., Chen H., Cheuk R.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Kim C., Lam B., Lin J., Miranda M., Narusaka M., Nguyen M.,
RA Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A.,
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB008268; BAB09859.1; -; Genomic_DNA.
DR EMBL; AY045893; AAK76567.1; -; mRNA.
DR EMBL; AY133825; AAM91759.1; -; rRNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006814; P:sodium ion transport; IEA.
DR InterPro; IPR001898; Na/sul_sympo.
DR Pfam; PF00939; Na sulph sym; 1.
DR TIGRFAMs; TIGR00785; dss; 1.
```

```
SQ SEQUENCE 549 AA; 58747 MW; 61FD036CDB0D45F4 CRC64;

Query Match          93.3%; Score 42; DB 2; Length 549;
Best Local Similarity 85.7%; Pred. No. 1.4e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 66 KPQPPPP 72

RESULT 49
Q4RP94_TETNG PRELIMINARY; PRT; 579 AA.
AC Q4RP94;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 1 SCAF15008, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00031214001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontoidea; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Athouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAE01015008; CAG09788.1; -; Genomic_DNA.
FT NON_TER 1
FT NON_TER 579
SQ SEQUENCE 579 AA; 65337 MW; E861791CB94FE00A CRC64;

Query Match          93.3%; Score 42; DB 2; Length 579;
Best Local Similarity 85.7%; Pred. No. 1.5e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db 140 RPEPPPP 146

RESULT 50
Q6AHW2_PNECA PRELIMINARY; PRT; 619 AA.
AC Q6AHW2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Protease-1 (PRT1), putative (Fragment).
```

GN Name=PC11A11.01;
OS Pneumocystis carinii.
OC Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes; Pneumocystidaceae;
OC Pneumocystis.
OX NCBI_TaxID=4754;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15496914; DOI=10.1038/nature03025;
RA Jaillon O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castellani V., Katinka M., Vacherie B.,
RA Blimont C., Skalli Z., Cattolico L., Poulain J., Berardinis Vd.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.-P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.-N., Guigs R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quittier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Crollius H.R.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Renauld H., Keely S.P., Barrell B., Stringer J.R., Berriman M.,
RA Aletti M.A., Hall N.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR716157; CAH17853.1; -; Genomic DNA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0005199; F:structural constituent of cell wall; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002029; Pept_S8_S53.
DR InterPro; IPR003882; Pstl_extensin.
DR Pfam; PF01483; P:protease; 1.
DR Pfam; PF00082; Peptidase S8; 1.
DR PRINTS; PR01218; PSTIEXTENSIN.
DR ProDom; PD000717; P:protnconvertsP; 1.
KW Protease.
FT NON TER 1 1
SQ SEQUENCE 619 AA; 66355 MW; 11601A87528508D3 CRC64;

Query Match 93.3%; Score 42; DB 2; Length 619;
Best Local Similarity 85.7%; Pred. No. 1.6e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPQPPPP 7
Db :|||||
528 KPQPPPP 534

Search completed: April 6, 2006, 09:39:50
Job time : 109.105 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.
OM protein - protein search, using sw model
Run on: April 6, 2006, 09:22:55 ; Search time 134.211 Seconds
(without alignments)
39.286 Million cell updates/sec

Title: US-10-632-388-307
Perfect score: 48
Sequence: 1 XXXRPLPLXP 12
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 2443163 seqs, 439378781 residues
Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : A Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				% Query		Description	
Result No.	Score	Match	Length	DB	ID		
1	48	100.0	12	2	AAR93346	Aar93346 SRC/LYN p	
2	48	100.0	12	2	AAR93356	Aar93356 FYN/LYN p	
3	48	100.0	15	2	AAR93471	Aar93471 GST-SRC p	
4	47	97.9	11	3	AAB21130	Aab21130 Src homol	
5	47	97.9	11	3	AAB21125	Aab21125 Src homol	
6	47	97.9	12	2	AAR93351	Aar93351 FYN prote	
7	47	97.9	12	2	AAR93359	Aar93359 LYN prote	
8	47	97.9	12	2	AAR93344	Aar93344 SRC prote	
9	47	97.9	12	3	AAB17254	Aab17254 SH3 antag	
10	47	97.9	12	3	AAB17251	Aab17251 SH3 antag	
11	47	97.9	12	5	ABB73244	Abb73244 Src homol	
12	47	97.9	12	5	ABB73247	Abb73247 Src homol	
13	47	97.9	12	7	ADJ73401	Adj73401 SH3 antag	
14	47	97.9	12	7	ADJ73398	Adj73398 SH3 antag	
15	47	97.9	12	8	ADJ53032	Adj53032 CH1 delet	
16	47	97.9	12	8	ADJ53035	Adj53035 CH1 delet	
17	47	97.9	12	8	ADJ51993	Adj51993 CH1 delet	
18	47	97.9	12	8	ADJ51996	Adj51996 CH1 delet	
19	47	97.9	15	2	AAR93461	Aar93461 GST-PI3K	
20	47	97.9	15	2	AAR93457	Aar93457 GST-LYN p	
21	47	97.9	15	2	AAR93460	Aar93460 GST-LYN p	
22	47	97.9	15	2	AAR93458	Aar93458 GST-LYN p	
23	47	97.9	15	2	AAR93459	Aar93459 GST-LYN p	
24	47	97.9	15	2	AAW05451	Aaw05451 SH3-bindi	

25	47	97.9	15	2	AAW37659	Aaw37659 PPPPY mot
26	47	97.9	15	2	AAW38924	Aaw38924 Peptide r
27	47	97.9	15	7	ADB49285	Adb49285 Novel WW
28	47	97.9	17	3	AAB17217	Aab17217 Mast cell
29	47	97.9	17	5	ABB73208	Abb73208 Mast cell
30	47	97.9	17	7	ADJ73362	Adj73362 Mast cell
31	47	97.9	17	8	ADJ52996	Adj52996 CH1 delet
32	47	97.9	17	8	ADJ51957	Adj51957 CH1 delet
33	47	97.9	22	2	AAy41645	Aay41645 Human pep
34	47	97.9	25	2	AAy41612	Aay41612 Mammalian
35	47	97.9	613	6	ABR42419	Abr42419 Human pot
36	46	95.8	12	2	AAR93362	Aar93362 LYN/PI3K
37	46	95.8	12	2	AAR93377	Aar93377 Grb-2 pro
38	46	95.8	12	2	AAR93345	Aar93345 SRC prote
39	46	95.8	12	3	AAB17255	Aab17255 SH3 antag
40	46	95.8	12	5	ABB73248	Abb73248 Src homol
41	46	95.8	12	7	ADJ73402	Adj73402 SH3 antag
42	46	95.8	12	8	ADJ53036	Adj53036 CH1 delet
43	46	95.8	12	8	ADJ51997	Adj51997 CH1 delet
44	46	95.8	14	2	AAR93386	Aar93386 GST-SRC S
45	46	95.8	14	2	AAR93404	Aar93404 GST-FYN S
46	46	95.8	14	2	AAR93450	Aar93450 GST-PI3K
47	46	95.8	14	2	AAR93388	Aar93388 GST-SRC S
48	46	95.8	14	2	AAR93392	Aar93392 GST-FYN S
49	46	95.8	14	2	AAR93400	Aar93400 GST-FYN S
50	46	95.8	14	2	AAR93428	Aar93428 GST-YES S
51	46	95.8	14	2	AAR93434	Aar93434 GST-YES S
52	46	95.8	14	2	AAR93439	Aar93439 GST-YES S
53	46	95.8	14	2	AAR93381	Aar93381 GST-SRC S
54	46	95.8	14	2	AAR93383	Aar93383 GST-SRC S
55	46	95.8	14	2	AAR93410	Aar93410 GST-LYN S
56	46	95.8	14	2	AAR93452	Aar93452 GST-PI3K
57	46	95.8	14	2	AAR93396	Aar93396 GST-FYN S
58	46	95.8	14	2	AAR93402	Aar93402 GST-FYN S
59	46	95.8	14	2	AAR93427	Aar93427 GST-YES S
60	46	95.8	14	2	AAR93432	Aar93432 GST-YES S
61	46	95.8	14	2	AAR93385	Aar93385 GST-YES S
62	46	95.8	14	2	AAR93419	Aar93419 GST-LYN S
63	46	95.8	14	2	AAR93420	Aar93420 GST-LYN S
64	46	95.8	14	2	AAR93431	Aar93431 GST-YES S
65	46	95.8	14	2	AAR93446	Aar93446 GST-PI3K
66	46	95.8	14	2	AAR93390	Aar93390 GST-SRC S
67	46	95.8	14	2	AAR93399	Aar93399 GST-FYN S
68	46	95.8	14	2	AAR93406	Aar93406 GST-FYN S
69	46	95.8	14	2	AAR93411	Aar93411 GST-LYN S
70	46	95.8	14	2	AAR93451	Aar93451 GST-PI3K
71	46	95.8	14	2	AAR93455	Aar93455 GST-PI3K
72	46	95.8	14	2	AAR93401	Aar93401 GST-FYN S
73	46	95.8	14	2	AAR93414	Aar93414 GST-LYN S
74	46	95.8	14	2	AAR93418	Aar93418 GST-LYN S
75	46	95.8	14	2	AAR93389	Aar93389 GST-SRC S
76	46	95.8	14	2	AAR93405	Aar93405 GST-FYN S
77	46	95.8	14	2	AAR93433	Aar93433 GST-YES S
78	46	95.8	14	2	AAR93444	Aar93444 GST-PI3K
79	46	95.8	14	2	AAR93403	Aar93403 GST-FYN S
80	46	95.8	14	2	AAR93424	Aar93424 GST-LYN S
81	46	95.8	14	2	AAR93387	Aar93387 GST-SRC S
82	46	95.8	14	2	AAR93412	Aar93412 GST-LYN S
83	46	95.8	14	2	AAR93425	Aar93425 GST-YES S
84	46	95.8	14	2	AAR93441	Aar93441 GST-YES S
85	46	95.8	14	2	AAR93454	Aar93454 GST-PI3K
86	46	95.8	14	2	AAR93393	Aar93393 GST-FYN S
87	46	95.8	14	2	AAR93398	Aar93398 GST-FYN S
88	46	95.8	14	2	AAR93413	Aar93413 GST-LYN S
89	46	95.8	14	2	AAR93416	Aar93416 GST-LYN S
90	46	95.8	14	2	AAR93422	Aar93422 GST-LYN S
91	46	95.8	14	2	AAR93436	Aar93436 GST-YES S
92	46	95.8	14	2	AAR93442	Aar93442 GST-YES S
93	46	95.8	14	2	AAR93445	Aar93445 GST-PI3K
94	46	95.8	14	2	AAR93391	Aar93391 GST-FYN S
95	46	95.8	14	2	AAR93394	Aar93394 GST-FYN S
96	46	95.8	14	2	AAR93417	Aar93417 GST-LYN S
97	46	95.8	14	2	AAR93429	Aar93429 GST-YES S

98 46 95.8 14 2 AAR933430 Aar933430 GST-YES S 171 45 93.8 85 6 ABO02918 Abo02918 Human sec
99 46 95.8 14 2 AAR933435 Aar933435 GST-YES S 172 45 93.8 85 6 ABR75072 Abr75072 Human sec
100 46 95.8 14 2 AAR933437 Aar933437 GST-YES S 173 45 93.8 85 6 ABR94834 Abr94834 Human sec
101 46 95.8 14 2 AAR933438 Aar933438 GST-YES S 174 45 93.8 85 6 ABO25164 Abo25164 Novel hum
102 46 95.8 14 2 AAR933440 Aar933440 GST-YES S 175 45 93.8 85 6 ABO33760 Abo33760 Novel hum
103 46 95.8 14 2 AAR933397 Aar933397 GST-FYN S 176 45 93.8 85 6 ABU85807 Abu85807 Human PRO
104 46 95.8 14 2 AAR933407 Aar933407 GST-FYN S 177 45 93.8 85 6 ABU98967 Abu98967 Novel hum
105 46 95.8 14 2 AAR933447 Aar933447 GST-PI3K 178 45 93.8 85 6 ABU98182 Abu98182 Novel hum
106 46 95.8 14 2 AAR933453 Aar933453 GST-PI3K 179 45 93.8 85 6 ABU91888 Abu91888 Novel hum
107 46 95.8 14 2 AAR933456 Aar933456 GST-PI3K 180 45 93.8 85 6 ABU89581 Abu89581 Human PRO
108 46 95.8 14 2 AAR933426 Aar933426 GST-YES S 181 45 93.8 85 6 ABU86422 Abu86422 Human sec
109 46 95.8 14 2 AAR933443 Aar933443 GST-YES S 182 45 93.8 85 6 ABU67635 Abu67635 Human sec
110 46 95.8 14 2 AAR933448 Aar933448 GST-PI3K 183 45 93.8 85 6 ABU80663 Abu80663 Human PRO
111 46 95.8 14 2 AAR933382 Aar933382 GST-SRC S 184 45 93.8 85 6 ABR99581 Abr99581 Human sec
112 46 95.8 14 2 AAR933384 Aar933384 GST-SRC S 185 45 93.8 85 6 ABR98971 Abr98971 Human sec
113 46 95.8 14 2 AAR933408 Aar933408 GST-LYN S 186 45 93.8 85 6 ABO16494 Abo16494 Human sec
114 46 95.8 14 2 AAR933409 Aar933409 GST-LYN S 187 45 93.8 85 6 ABR92394 Abr92394 Human sec
115 46 95.8 14 2 AAR933421 Aar933421 GST-LYN S 188 45 93.8 85 6 ABO19035 Abo19035 Human sec
116 46 95.8 14 2 AAR933449 Aar933449 GST-PI3K 189 45 93.8 85 6 ABR78456 Abr78456 Human sec
117 46 95.8 14 2 AAR933395 Aar933395 GST-FYN S 190 45 93.8 85 6 ABU85192 Abu85192 Novel hum
118 46 95.8 14 2 AAR933415 Aar933415 GST-LYN S 191 45 93.8 85 6 ABO00331 Abo00331 Novel hum
119 46 95.8 14 2 AAR933423 Aar933423 GST-LYN S 192 45 93.8 85 6 ABO11663 Abo11663 Human sec
120 46 95.8 282 4 AAG62649 Aag62649 Porcine F 193 45 93.8 85 6 ABO02308 Abo02308 Human sec
121 45 93.8 45 2 AAY36531 Aay36531 Fragment 194 45 93.8 85 6 ADA40702 Ada40702 Human sec
122 45 93.8 45 6 ADA12079 Ada12079 Human nov 195 45 93.8 85 6 ABU88882 Abu88882 Novel hum
123 45 93.8 68 4 AAB51247 Aab51247 Human ami 196 45 93.8 85 6 ABU67282 Abu67282 Novel hum
124 45 93.8 78 2 AAY12633 Aay12633 Human 5' 197 45 93.8 85 6 ABU83577 Abu83577 Human sec
125 45 93.8 83 2 AAY12632 Aay12632 Human 5' 198 45 93.8 85 6 ABO06378 Abo06378 Novel hum
126 45 93.8 83 4 AAB50241 Aab50241 Human bre 199 45 93.8 85 6 ABR59414 Abr59414 Human sec
127 45 93.8 83 4 AAU33342 Aau33342 Human bre 200 45 93.8 85 6 ABO09476 Abo09476 Human sec
128 45 93.8 83 5 ABG66759 Abg66759 Human pro 201 45 93.8 85 6 ABO19340 Abo19340 Novel hum
129 45 93.8 83 5 AAO15776 Aao15776 Human bre 202 45 93.8 85 6 ABO11358 Abo11358 Human sec
130 45 93.8 83 5 ABG78909 Abg78909 Human bre 203 45 93.8 85 6 ABR66976 Abr66976 Human sec
131 45 93.8 83 6 ABJ37732 Abj37732 Human tum 204 45 93.8 85 6 ABO16189 Abo16189 Human sec
132 45 93.8 83 7 ADL92724 Adl92724 Human bre 205 45 93.8 85 6 ABO13895 Abo13895 Human sec
133 45 93.8 83 8 ADE44014 Ade44014 Human bre 206 45 93.8 85 6 ABU65798 Abu65798 Human sec
134 45 93.8 85 2 AAW59652 Aaw59652 Amino aci 207 45 93.8 85 6 ABO07646 Abo07646 Human PRO
135 45 93.8 85 2 AAY04146 Aay04146 Human 5' 208 45 93.8 85 6 ABO03833 Abo03833 Human sec
136 45 93.8 85 2 AAW93615 Aaw93615 Human 5' 209 45 93.8 85 6 ABR67281 Abr67281 Human sec
137 45 93.8 85 2 AAY04164 Aay04164 Human 5' 210 45 93.8 85 6 ABO15884 Abo15884 Human sec
138 45 93.8 85 2 AAY36038 Aay36038 Extended 211 45 93.8 85 6 ABU56165 Abu56165 Human sec
139 45 93.8 85 2 AAY36037 Aay36037 Extended 212 45 93.8 85 6 ABU65493 Abu65493 Human PRO
140 45 93.8 85 2 AAY35880 Aay35880 Extended 213 45 93.8 85 6 ABU95438 Abu95438 Novel hum
141 45 93.8 85 2 AAY59692 Aay59692 Secreted 214 45 93.8 85 6 ABU71341 Abu71341 Human PRO
142 45 93.8 85 2 AAY59643 Aay59643 Secreted 215 45 93.8 85 6 ABO07951 Abo07951 Human PRO
143 45 93.8 85 2 AAY01589 Aay01589 Secreted 216 45 93.8 85 6 ABR70192 Abr70192 Human sec
144 45 93.8 85 2 AAY12981 Aay12981 Human sec 217 45 93.8 85 6 ABR69525 Abr69525 Human sec
145 45 93.8 85 2 AAY25449 Aay25449 Human sec 218 45 93.8 85 6 ABO01666 Abo01666 Human PRO
146 45 93.8 85 2 AAY12670 Aay12670 Human 5' 219 45 93.8 85 6 ABU81468 Abu81468 Human PRO
147 45 93.8 85 3 AAG00012 Aag00012 Human sec 220 45 93.8 85 6 ABR60265 Abr60265 Human sec
148 45 93.8 85 4 AAB31193 Aab31193 Amino aci 221 45 93.8 85 6 ABR68000 Abr68000 Human sec
149 45 93.8 85 4 AAU29253 Aau29253 Human PRO 222 45 93.8 85 6 ABR65388 Abr65388 Human sec
150 45 93.8 85 4 AAB51245 Aab51245 Human ami 223 45 93.8 85 6 ABR68610 Abr68610 Human sec
151 45 93.8 85 4 AAB90699 Aab90699 Human CH3 224 45 93.8 85 6 ABR72022 Abr72022 Human sec
152 45 93.8 85 5 AAU83647 Aau83647 Human PRO 225 45 93.8 85 6 ABU85502 Abu85502 Human PRO
153 45 93.8 85 5 ABB84934 Abb84934 Human PRO 226 45 93.8 85 6 ABU89192 Abu89192 Human sec
154 45 93.8 85 5 AAE20461 Aae20461 Human tum 227 45 93.8 85 6 ABU83272 Abu83272 Human sec
155 45 93.8 85 5 ABB95540 Abb95540 Human ang 228 45 93.8 85 6 ABU95128 Abu95128 Novel hum
156 45 93.8 85 5 ADY31852 Ady31852 Novel hum 229 45 93.8 85 6 ABU90676 Abu90676 Novel hum
157 45 93.8 85 6 ABU58629 Abu58629 Human PRO 230 45 93.8 85 6 ABU84187 Abu84187 Human sec
158 45 93.8 85 6 ABU88177 Abu88177 Novel hum 231 45 93.8 85 6 ABU93838 Abu93838 Novel hum
159 45 93.8 85 6 ABR66366 Abr66366 Human sec 232 45 93.8 85 6 ABR65083 Abr65083 Human sec
160 45 93.8 85 6 ABR65756 Abr65756 Human sec 233 45 93.8 85 6 ABR68915 Abr68915 Human sec
161 45 93.8 85 6 ABR65756 Abr65756 Human sec 234 45 93.8 85 6 ABO06731 Abo06731 Human sec
162 45 93.8 85 6 ABU99696 Abu99696 Human sec 235 45 93.8 85 6 ABR99276 Abr99276 Human sec
163 45 93.8 85 6 ABU82935 Abu82935 Human PRO 236 45 93.8 85 6 ABU57160 Abu57160 Human PRO
164 45 93.8 85 6 ABU90056 Abu90056 Novel hum 237 45 93.8 85 6 ABU72050 Abu72050 Novel hum
165 45 93.8 85 6 ABR68305 Abr68305 Human sec 238 45 93.8 85 6 ABU86112 Abu86112 Novel hum
166 45 93.8 85 6 ADA56852 Ada56852 Human sec 239 45 93.8 85 6 ABU82399 Abu82399 Novel hum
167 45 93.8 85 6 ABU96358 Abu96358 Novel hum 240 45 93.8 85 6 ABU87410 Abu87410 Human PRO
168 45 93.8 85 6 ABU92789 Abu92789 Human sec 241 45 93.8 85 6 ABU67151 Abu67151 Novel hum
169 45 93.8 85 6 ABU80794 Abu80794 Human PRO 242 45 93.8 85 6 ABU83882 Abu83882 Human sec
170 45 93.8 85 6 ABO08866 Abo08866 Human sec 243 45 93.8 85 6 ABO08256 Abo08256 Human PRO

171 45 93.8 85 6 ABO02918 Abo02918 Human sec
172 45 93.8 85 6 ABR75072 Abr75072 Human sec
173 45 93.8 85 6 ABR94834 Abr94834 Human sec
174 45 93.8 85 6 ABO25164 Abo25164 Novel hum
175 45 93.8 85 6 ABO33760 Abo33760 Novel hum
176 45 93.8 85 6 ABU85807 Abu85807 Human PRO
177 45 93.8 85 6 ABU98967 Abu98967 Novel hum
178 45 93.8 85 6 ABU98182 Abu98182 Novel hum
179 45 93.8 85 6 ABU91888 Abu91888 Novel hum
180 45 93.8 85 6 ABU89581 Abu89581 Human PRO
181 45 93.8 85 6 ABU86422 Abu86422 Human sec
182 45 93.8 85 6 ABU67635 Abu67635 Human sec
183 45 93.8 85 6 ABU80663 Abu80663 Human PRO
184 45 93.8 85 6 ABR99581 Abr99581 Human sec
185 45 93.8 85 6 ABR98971 Abr98971 Human sec
186 45 93.8 85 6 ABO16494 Abo16494 Human sec
187 45 93.8 85 6 ABR92394 Abr92394 Human sec
188 45 93.8 85 6 ABO19035 Abo19035 Human sec
189 45 93.8 85 6 ABR78456 Abr78456 Human sec
190 45 93.8 85 6 ABU85192 Abu85192 Novel hum
191 45 93.8 85 6 ABO00331 Abo00331 Novel hum
192 45 93.8 85 6 ABO11663 Abo11663 Human sec
193 45 93.8 85 6 ABO02308 Abo02308 Human sec
194 45 93.8 85 6 ADA40702 Ada40702 Human sec
195 45 93.8 85 6 ABU88882 Abu88882 Novel hum
196 45 93.8 85 6 ABU67282 Abu67282 Novel hum
197 45 93.8 85 6 ABU83577 Abu83577 Human sec
198 45 93.8 85 6 ABO06378 Abo06378 Novel hum
199 45 93.8 85 6 ABR59414 Abr59414 Human sec
200 45 93.8 85 6 ABO09476 Abo09476 Human sec
201 45 93.8 85 6 ABO19340 Abo19340 Novel hum
202 45 93.8 85 6 ABO11358 Abo11358 Human sec
203 45 93.8 85 6 ABR66976 Abr66976 Human sec
204 45 93.8 85 6 ABO16189 Abo16189 Human sec
205 45 93.8 85 6 ABO13895 Abo13895 Human sec
206 45 93.8 85 6 ABU65798 Abu65798 Human sec
207 45 93.8 85 6 ABO07646 Abo07646 Human PRO
208 45 93.8 85 6 ABO03833 Abo03833 Human sec
209 45 93.8 85 6 ABR67281 Abr67281 Human sec
210 45 93.8 85 6 ABO15884 Abo15884 Human sec
211 45 93.8 85 6 ABU56165 Abu56165 Human sec
212 45 93.8 85 6 ABU65493 Abu65493 Human PRO
213 45 93.8 85 6 ABU95438 Abu95438 Novel hum
214 45 93.8 85 6 ABU71341 Abu71341 Human PRO
215 45 93.8 85 6 ABO07951 Abo07951 Human PRO
216 45 93.8 85 6 ABR70192 Abr70192 Human sec
217 45 93.8 85 6 ABR69525 Abr69525 Human sec
218 45 93.8 85 6 ABO01666 Abo01666 Human PRO
219 45 93.8 85 6 ABU81468 Abu81468 Human PRO
220 45 93.8 85 6 ABR60265 Abr60265 Human sec
221 45 93.8 85 6 ABR68000 Abr68000 Human sec
222 45 93.8 85 6 ABR65388 Abr65388 Human sec
223 45 93.8 85 6 ABR68610 Abr68610 Human sec
224 45 93.8 85 6 ABR72022 Abr72022 Human sec
225 45 93.8 85 6 ABU85502 Abu85502 Human PRO
226 45 93.8 85 6 ABU89192 Abu89192 Human sec
227 45 93.8 85 6 ABU83272 Abu83272 Human sec
228 45 93.8 85 6 ABU95128 Abu95128 Novel hum
229 45 93.8 85 6 ABU90676 Abu90676 Novel hum
230 45 93.8 85 6 ABU84187 Abu84187 Human sec
231 45 93.8 85 6 ABU93838 Abu93838 Novel hum
232 45 93.8 85 6 ABR65083 Abr65083 Human sec
233 45 93.8 85 6 ABR68915 Abr68915 Human sec
234 45 93.8 85 6 ABO06731 Abo06731 Human sec
235 45 93.8 85 6 ABR99276 Abr99276 Human sec
236 45 93.8 85 6 ABU57160 Abu57160 Human PRO
237 45 93.8 85 6 ABU72050 Abu72050 Novel hum
238 45 93.8 85 6 ABU86112 Abu86112 Novel hum
239 45 93.8 85 6 ABU82399 Abu82399 Novel hum
240 45 93.8 85 6 ABU87410 Abu87410 Human PRO
241 45 93.8 85 6 ABU67151 Abu67151 Novel hum
242 45 93.8 85 6 ABU83882 Abu83882 Human sec
243 45 93.8 85 6 ABO08256 Abo08256 Human PRO

Abo02918 Human sec
Abr75072 Human sec
Abr94834 Human sec
Abo25164 Novel hum
Abo33760 Novel hum
Abu85807 Human PRO
Abu98967 Novel hum
Abu98182 Novel hum
Abu91888 Novel hum
Abu89581 Human PRO
Abu86422 Human sec
Abu67635 Human sec
Abu80663 Human PRO
Abr99581 Human sec
Abr98971 Human sec
Abo16494 Human sec
Abr92394 Human sec
Abo19035 Human sec
Abr78456 Human sec
Abu85192 Novel hum
Abo00331 Novel hum
Abo11663 Human sec
Abo02308 Human sec
Ada40702 Human sec
Abu88882 Novel hum
Abu67282 Novel hum
Abu83577 Human sec
Abo06378 Novel hum
Abr59414 Human sec
Abo09476 Human sec
Abo19340 Novel hum
Abo11358 Human sec
Abr66976 Human sec
Abo16189 Human sec
Abo13895 Human sec
Abu65798 Human sec
Abo07646 Human PRO
Abo03833 Human sec
Abr67281 Human sec
Abo15884 Human sec
Abu56165 Human sec
Abu65493 Human PRO
Abu95438 Novel hum
Abu71341 Human PRO
Abo07951 Human PRO
Abr70192 Human sec
Abr69525 Human sec
Abo01666 Human PRO
Abu81468 Human PRO
Abr60265 Human sec
Abr68000 Human sec
Abr65388 Human sec
Abr68610 Human sec
Abr72022 Human sec
Abu85502 Human PRO
Abu89192 Human sec
Abu83272 Human sec
Abu95128 Novel hum
Abu90676 Novel hum
Abu84187 Human sec
Abu93838 Novel hum
Abr65083 Human sec
Abr68915 Human sec
Abo06731 Human sec
Abr99276 Human sec
Abu57160 Human PRO
Abu72050 Novel hum
Abu86112 Novel hum
Abu82399 Novel hum
Abu87410 Human PRO
Abu67151 Novel hum
Abu83882 Human sec
Abo08256 Human PRO

244	45	93.8	85	6	ABU81967	Novel hum	317	45	93.8	85	6	ABO48262	Human sec
245	45	93.8	85	6	ABU66131	Novel hum	318	45	93.8	85	6	ABR93004	Human sec
246	45	93.8	85	6	ABR59960	Human sec	319	45	93.8	85	6	ABO24765	Human sec
247	45	93.8	85	6	ABU94148	Novel hum	320	45	93.8	85	6	ABM11776	Human sec
248	45	93.8	85	6	ABO00021	Novel hum	321	45	93.8	85	6	ABM02877	Human sec
249	45	93.8	85	6	ABU82103	Novel hum	322	45	93.8	85	6	ABM16173	Human sec
250	45	93.8	85	6	ABR66671	Human sec	323	45	93.8	85	6	ABO27734	Human sec
251	45	93.8	85	6	ABR91089	Human sec	324	45	93.8	85	6	ABM29225	Human sec
252	45	93.8	85	6	ABG75947	Human ant	325	45	93.8	85	6	ABM07201	Human sec
253	45	93.8	85	6	ABU94516	Human PRO	326	45	93.8	85	6	ABM21295	Human sec
254	45	93.8	85	6	ABU79398	Human PRO	327	45	93.8	85	6	ABM09641	Human sec
255	45	93.8	85	6	ABU86727	Human sec	328	45	93.8	85	6	ABO41511	Human sec
256	45	93.8	85	6	ABU87032	Novel hum	329	45	93.8	85	6	ABO36326	Human PRO
257	45	93.8	85	6	ABU94821	Human PRO	330	45	93.8	85	6	ABO43855	Human PRO
258	45	93.8	85	6	ABO04748	Human PRO	331	45	93.8	85	6	ABM76555	Human sec
259	45	93.8	85	6	ABR70497	Human sec	332	45	93.8	85	6	ABM76251	Human sec
260	45	93.8	85	6	ABU98662	Human PRO	333	45	93.8	85	6	ABM25870	Human sec
261	45	93.8	85	6	ABR66061	Human sec	334	45	93.8	85	6	ABM26175	Human sec
262	45	93.8	85	6	ABR64778	Human sec	335	45	93.8	85	6	ABO03528	Human sec
263	45	93.8	85	6	ABU79703	Human PRO	336	45	93.8	85	6	ABO02613	Human sec
264	45	93.8	85	6	ABU93094	Human sec	337	45	93.8	85	6	ABR90784	Human sec
265	45	93.8	85	6	ABU96053	Human PRO	338	45	93.8	85	6	ABR73852	Human sec
266	45	93.8	85	6	ABU91273	Novel hum	339	45	93.8	85	6	ABO17104	Human sec
267	45	93.8	85	6	ABU90366	Novel hum	340	45	93.8	85	6	ABR94529	Human sec
268	45	93.8	85	6	ABO09781	Human sec	341	45	93.8	85	6	ABR76036	Human sec
269	45	93.8	85	6	ABO11053	Human sec	342	45	93.8	85	6	ABR71412	Human sec
270	45	93.8	85	6	ABR71107	Human sec	343	45	93.8	85	6	ABR93309	Human sec
271	45	93.8	85	6	ABU87715	Human PRO	344	45	93.8	85	6	ABR93614	Human sec
272	45	93.8	85	6	ABU91583	Human PRO	345	45	93.8	85	6	ABR88039	Human sec
273	45	93.8	85	6	ABU84797	Human sec	346	45	93.8	85	6	ABO28039	Human sec
274	45	93.8	85	6	ABR69887	Human sec	347	45	93.8	85	6	ABO30174	Human sec
275	45	93.8	85	6	ABU80264	Human PRO	348	45	93.8	85	6	ABO33383	Human PRO
276	45	93.8	85	6	ABU93533	Human PRO	349	45	93.8	85	6	ABM05071	Human sec
277	45	93.8	85	6	ABO10086	Human sec	350	45	93.8	85	6	ABM09031	Human sec
278	45	93.8	85	6	ABO09171	Human sec	351	45	93.8	85	6	ABO36631	Human sec
279	45	93.8	85	6	ABU10739	Human sec	352	45	93.8	85	6	ABO35716	Human sec
280	45	93.8	85	6	ABU95748	Human PRO	353	45	93.8	85	6	ABO39681	Human sec
281	45	93.8	85	6	ABU79793	Human sec	354	45	93.8	85	6	ABM10556	Human sec
282	45	93.8	85	6	ABU96957	Novel hum	355	45	93.8	85	6	ABM12081	Human sec
283	45	93.8	85	6	ABR70802	Human sec	356	45	93.8	85	6	ABO52227	Human PRO
284	45	93.8	85	6	ABO05153	Novel hum	357	45	93.8	85	6	ABO52532	Human PRO
285	45	93.8	85	6	ABO08561	Human sec	358	45	93.8	85	6	ABO23850	Human sec
286	45	93.8	85	6	ABO05768	Human sec	359	45	93.8	85	6	ABR97336	Human sec
287	45	93.8	85	6	ABR74157	Human sec	360	45	93.8	85	6	ABR87124	Human sec
288	45	93.8	85	6	ABR95749	Human sec	361	45	93.8	85	6	ABM11166	Human sec
289	45	93.8	85	6	ABR81046	Human sec	362	45	93.8	85	6	ABM28310	Human sec
290	45	93.8	85	6	ABR81351	Human sec	363	45	93.8	85	6	ABO32309	Human sec
291	45	93.8	85	6	ABM01047	Human sec	364	45	93.8	85	6	ABM15436	Human sec
292	45	93.8	85	6	ABR88649	Human sec	365	45	93.8	85	6	ABM06591	Human sec
293	45	93.8	85	6	ABM77470	Human sec	366	45	93.8	85	6	ABM04402	Human sec
294	45	93.8	85	6	ABO28954	Human sec	367	45	93.8	85	6	ABM22515	Human sec
295	45	93.8	85	6	ABO31699	Human sec	368	45	93.8	85	6	ABM07811	Human sec
296	45	93.8	85	6	ABM08116	Human sec	369	45	93.8	85	6	ABO40901	Human sec
297	45	93.8	85	6	ABO40596	Human sec	370	45	93.8	85	6	ABM35548	Human sec
298	45	93.8	85	6	ABO36021	Human PRO	371	45	93.8	85	6	ABM33311	Human sec
299	45	93.8	85	6	ABO44160	Human PRO	372	45	93.8	85	6	ABO52837	Human PRO
300	45	93.8	85	6	ADA78212	Human sec	373	45	93.8	85	6	ABO50397	Human sec
301	45	93.8	85	6	ABM24955	Human sec	374	45	93.8	85	6	ABU99391	Human sec
302	45	93.8	85	6	ABO03223	Human sec	375	45	93.8	85	6	ABO04443	Human sec
303	45	93.8	85	6	ABR90479	Human sec	376	45	93.8	85	6	ABO06073	Human sec
304	45	93.8	85	6	ABM17393	Human sec	377	45	93.8	85	6	ABM18613	Human sec
305	45	93.8	85	6	ABR95139	Human sec	378	45	93.8	85	6	ABR97641	Human sec
306	45	93.8	85	6	ABR95444	Human sec	379	45	93.8	85	6	ABR80741	Human sec
307	45	93.8	85	6	ABJ72283	Human PRO	380	45	93.8	85	6	ABM01352	Human sec
308	45	93.8	85	6	ABO21682	Human sec	381	45	93.8	85	6	ABR88954	Human sec
309	45	93.8	85	6	ABR97946	Human sec	382	45	93.8	85	6	ABM13606	Human sec
310	45	93.8	85	6	ABR87734	Human sec	383	45	93.8	85	6	ABM20990	Human sec
311	45	93.8	85	6	ABM77775	Human sec	384	45	93.8	85	6	ABO42121	Human sec
312	45	93.8	85	6	ABM28005	Human sec	385	45	93.8	85	6	ABO42731	Human sec
313	45	93.8	85	6	ABM06286	Human sec	386	45	93.8	85	6	ABM10251	Human sec
314	45	93.8	85	6	ABM03792	Human sec	387	45	93.8	85	6	ABO38766	Human sec
315	45	93.8	85	6	ABM35243	Human sec	388	45	93.8	85	6	ABM33006	Human sec
316	45	93.8	85	6	ABM26480	Human sec	389	45	93.8	85	6	ABM22820	Human sec

390	45	93.8	85	6	ABM75031	Abm75031 Human sec	463	45	93.8	85	6	ABM25565	Abm25565 Human sec
391	45	93.8	85	6	ADA80004	Ada80004 Human sec	464	45	93.8	85	6	ABM20075	Abm20075 Human sec
392	45	93.8	85	6	ABR96421	Abm96421 Human sec	465	45	93.8	85	6	ABO46981	Abm46981 Human PRO
393	45	93.8	85	6	ABM02572	Abm02572 Human sec	466	45	93.8	85	6	ABO47286	Abm47286 Human PRO
394	45	93.8	85	6	ABR86514	Abm86514 Human sec	467	45	93.8	85	6	ADA83529	Ada83529 Human sec
395	45	93.8	85	6	ABR86819	Abm86819 Human sec	468	45	93.8	85	6	ABR71717	Abm71717 Human sec
396	45	93.8	85	6	ABM16783	Abm16783 Human sec	469	45	93.8	85	6	ABR72327	Abm72327 Human sec
397	45	93.8	85	6	ABM29835	Abm29835 Human sec	470	45	93.8	85	6	ABR98666	Abm98666 Human sec
398	45	93.8	85	6	ABO29259	Abm29259 Human sec	471	45	93.8	85	6	ABO07036	Abm07036 Human sec
399	45	93.8	85	6	ABM24040	Abm24040 Human sec	472	45	93.8	85	6	ABR84989	Abm84989 Human sec
400	45	93.8	85	6	ABM23430	Abm23430 Human sec	473	45	93.8	85	6	ABR73547	Abm73547 Human sec
401	45	93.8	85	6	ADA47228	Ada47228 Human sec	474	45	93.8	85	6	ABR76641	Abm76641 Human sec
402	45	93.8	85	6	ABM22210	Abm22210 Human sec	475	45	93.8	85	6	ABR73242	Abm73242 Human sec
403	45	93.8	85	6	ABO37851	Abm37851 Human sec	476	45	93.8	85	6	ABJ72411	Abj72411 Human PRO
404	45	93.8	85	6	ABM28615	Abm28615 Human sec	477	45	93.8	85	6	ABM18308	Abm18308 Human sec
405	45	93.8	85	6	ABM28920	Abm28920 Human sec	478	45	93.8	85	6	ABO20767	Abm20767 Human sec
406	45	93.8	85	6	ABM66564	Abm66564 Human sec	479	45	93.8	85	6	ABO25510	Abm25510 Human PRO
407	45	93.8	85	6	ABM75946	Abm75946 Human sec	480	45	93.8	85	6	ABO25815	Abm25815 Human PRO
408	45	93.8	85	6	ABM34226	Abm34226 Human sec	481	45	93.8	85	6	ABR94224	Abm94224 Human sec
409	45	93.8	85	6	ABM34531	Abm34531 Human sec	482	45	93.8	85	6	ABR80131	Abm80131 Human sec
410	45	93.8	85	6	ABO20462	Abm20462 Human sec	483	45	93.8	85	6	ABM11471	Abm11471 Human sec
411	45	93.8	85	6	ABO21377	Abm21377 Human sec	484	45	93.8	85	6	ABO33078	Abm33078 Human PRO
412	45	93.8	85	6	ABO22292	Abm22292 Human sec	485	45	93.8	85	6	ABO30784	Abm30784 Human sec
413	45	93.8	85	6	ABR96726	Abm96726 Human sec	486	45	93.8	85	6	ABO31089	Abm31089 Human sec
414	45	93.8	85	6	ABR85904	Abm85904 Human sec	487	45	93.8	85	6	ABM27395	Abm27395 Human sec
415	45	93.8	85	6	ABR99886	Abm99886 Human sec	488	45	93.8	85	6	ABM30140	Abm30140 Human sec
416	45	93.8	85	6	ABM00742	Abm00742 Human sec	489	45	93.8	85	6	ABM05676	Abm05676 Human sec
417	45	93.8	85	6	ABM00437	Abm00437 Human sec	490	45	93.8	85	6	ABM15741	Abm15741 Human sec
418	45	93.8	85	6	ABO29869	Abm29869 Human sec	491	45	93.8	85	6	ABM08726	Abm08726 Human sec
419	45	93.8	85	6	ABM23735	Abm23735 Human sec	492	45	93.8	85	6	ABO42426	Abm42426 Human sec
420	45	93.8	85	6	ABM29530	Abm29530 Human sec	493	45	93.8	85	6	ABO38156	Abm38156 Human sec
421	45	93.8	85	6	ABO38461	Abm38461 Human sec	494	45	93.8	85	6	ABO46066	Abm46066 Human PRO
422	45	93.8	85	6	ABO45761	Abm45761 Human PRO	495	45	93.8	85	6	ABM66869	Abm66869 Human sec
423	45	93.8	85	6	ABM20685	Abm20685 Human sec	496	45	93.8	85	6	ADB20572	Abm20572 Human sec
424	45	93.8	85	6	ADA81731	Ada81731 Human sec	497	45	93.8	85	6	ABM19770	Abm19770 Human sec
425	45	93.8	85	6	ABO16799	Abm16799 Human sec	498	45	93.8	85	6	ABO49482	Abm49482 Human sec
426	45	93.8	85	6	ABO18425	Abm18425 Human sec	499	45	93.8	85	6	ABO49787	Abm49787 Human sec
427	45	93.8	85	6	ABO22852	Abm22852 Human PRO	500	45	93.8	85	6	ADA78824	Ada78824 Human sec
428	45	93.8	85	6	ABO23157	Abm23157 Human PRO	501	45	93.8	85	6	ABO34306	Abm34306 Human sec
429	45	93.8	85	6	ABR92699	Abm92699 Human sec	502	45	93.8	85	6	ABR88344	Abm88344 Human sec
430	45	93.8	85	6	ABR81656	Abm81656 Human sec	503	45	93.8	85	6	ABM27090	Abm27090 Human sec
431	45	93.8	85	6	ABM78080	Abm78080 Human sec	504	45	93.8	85	6	ABM03487	Abm03487 Human sec
432	45	93.8	85	6	ABR89869	Abm89869 Human sec	505	45	93.8	85	6	ABO39986	Abm39986 Human sec
433	45	93.8	85	6	ABM26785	Abm26785 Human sec	506	45	93.8	85	7	ABO50092	Abm50092 Human sec
434	45	93.8	85	6	ABM13911	Abm13911 Human sec	507	45	93.8	85	7	ABO51007	Abm51007 Human sec
435	45	93.8	85	6	ABO28649	Abm28649 Human sec	508	45	93.8	85	7	ABO05463	Abm05463 Human sec
436	45	93.8	85	6	ABO30479	Abm30479 Human sec	509	45	93.8	85	7	ABR74767	Abm74767 Human sec
437	45	93.8	85	6	ABM07506	Abm07506 Human sec	510	45	93.8	85	7	ABR77246	Abm77246 Human sec
438	45	93.8	85	6	ABM04097	Abm04097 Human sec	511	45	93.8	85	7	ABM18003	Abm18003 Human sec
439	45	93.8	85	6	ABO37241	Abm37241 Human sec	512	45	93.8	85	7	ABR96054	Abm96054 Human sec
440	45	93.8	85	6	ABO41816	Abm41816 Human sec	513	45	93.8	85	7	ABO21987	Abm21987 Human sec
441	45	93.8	85	6	ABO35411	Abm35411 Human PRO	514	45	93.8	85	7	ABO20157	Abm20157 Human sec
442	45	93.8	85	6	ABM25260	Abm25260 Human sec	515	45	93.8	85	7	ABO24460	Abm24460 Human sec
443	45	93.8	85	6	ABO47652	Abm47652 Human sec	516	45	93.8	85	7	ABR86209	Abm86209 Human sec
444	45	93.8	85	6	ABO47957	Abm47957 Human sec	517	45	93.8	85	7	ABM10861	Abm10861 Human sec
445	45	93.8	85	6	ABO48567	Abm48567 Human sec	518	45	93.8	85	7	ABM76860	Abm76860 Human sec
446	45	93.8	85	6	ABO51617	Abm51617 Human PRO	519	45	93.8	85	7	ABR89564	Abm89564 Human sec
447	45	93.8	85	6	ABO51922	Abm51922 Human PRO	520	45	93.8	85	7	ABM12691	Abm12691 Human sec
448	45	93.8	85	6	ABO50702	Abm50702 Human sec	521	45	93.8	85	7	ABM05981	Abm05981 Human sec
449	45	93.8	85	6	ABR79826	Abm79826 Human sec	522	45	93.8	85	7	ABO35106	Abm35106 Human PRO
450	45	93.8	85	6	ABM17088	Abm17088 Human sec	523	45	93.8	85	7	ABM03182	Abm03182 Human sec
451	45	93.8	85	6	ABO18120	Abm18120 Human sec	524	45	93.8	85	7	ABM19160	Abm19160 Human sec
452	45	93.8	85	6	ABO21072	Abm21072 Human sec	525	45	93.8	85	7	ABM19465	Abm19465 Human sec
453	45	93.8	85	6	ABR97031	Abm97031 Human sec	526	45	93.8	85	7	ABO46676	Abm46676 Human PRO
454	45	93.8	85	6	ABM12386	Abm12386 Human sec	527	45	93.8	85	7	ABO49177	Abm49177 Human sec
455	45	93.8	85	6	ABM16478	Abm16478 Human sec	528	45	93.8	85	7	ABR69220	Abm69220 Human sec
456	45	93.8	85	6	ABM24345	Abm24345 Human sec	529	45	93.8	85	7	ABR89259	Abm89259 Human sec
457	45	93.8	85	6	ABM14826	Abm14826 Human sec	530	45	93.8	85	7	ABR72632	Abm72632 Human sec
458	45	93.8	85	6	ABM04707	Abm04707 Human sec	531	45	93.8	85	7	ABR74462	Abm74462 Human sec
459	45	93.8	85	6	ABM06896	Abm06896 Human sec	532	45	93.8	85	7	ABO18730	Abm18730 Human sec
460	45	93.8	85	6	ABM09336	Abm09336 Human sec	533	45	93.8	85	7	ABR80436	Abm80436 Human sec
461	45	93.8	85	6	ABO39376	Abm39376 Human sec	534	45	93.8	85	7	ABM01657	Abm01657 Human sec
462	45	93.8	85	6	ABM75641	Abm75641 Human sec	535	45	93.8	85	7	ABM02267	Abm02267 Human sec

536 45 93.8 85 7 ABR87429 Abr87429 Human sec
537 45 93.8 85 7 ABM12996 Abm12996 Human sec
538 45 93.8 85 7 ABM30750 Abm30750 Human sec
539 45 93.8 85 7 ABM24650 Abm24650 Human sec
540 45 93.8 85 7 ABO29564 Abo29564 Human sec
541 45 93.8 85 7 ABO31394 Abo31394 Human sec
542 45 93.8 85 7 ABM14521 Abm14521 Human sec
543 45 93.8 85 7 ABM09946 Abm09946 Human sec
544 45 93.8 85 7 ABO39071 Abo39071 Human sec
545 45 93.8 85 7 ABM34836 Abm34836 Human sec
546 45 93.8 85 7 ABO51312 Abo51312 Human sec
547 45 93.8 85 7 ABO04138 Abo04138 Human sec
548 45 93.8 85 7 ABO10608 Abo10608 Human PRO
549 45 93.8 85 7 ABR77851 Abr77851 Human sec
550 45 93.8 85 7 ABR79061 Abr79061 Human sec
551 45 93.8 85 7 ABO24155 Abo24155 Human sec
552 45 93.8 85 7 ABR93919 Abr93919 Human sec
553 45 93.8 85 7 ABM01962 Abm01962 Human sec
554 45 93.8 85 7 ABM78385 Abm78385 Human sec
555 45 93.8 85 7 ABR90174 Abr90174 Human sec
556 45 93.8 85 7 ABM27700 Abm27700 Human sec
557 45 93.8 85 7 ABM13301 Abm13301 Human sec
558 45 93.8 85 7 ABO32004 Abo32004 Human sec
559 45 93.8 85 7 ABM14216 Abm14216 Human sec
560 45 93.8 85 7 ABM08421 Abm08421 Human sec
561 45 93.8 85 7 ABO40291 Abo40291 Human sec
562 45 93.8 85 7 ABM74726 Abm74726 Human sec
563 45 93.8 85 7 ABM33921 Abm33921 Human sec
564 45 93.8 85 7 ABM20380 Abm20380 Human sec
565 45 93.8 85 7 ABO48872 Abo48872 Human sec
566 45 93.8 85 7 ABR72937 Abr72937 Human sec
567 45 93.8 85 7 ABO15579 Abo15579 Human sec
568 45 93.8 85 7 ABR85294 Abr85294 Human sec
569 45 93.8 85 7 ABO15274 Abo15274 Human sec
570 45 93.8 85 7 ABO17409 Abo17409 Human sec
571 45 93.8 85 7 ABO19851 Abo19851 Human sec
572 45 93.8 85 7 ABM17698 Abm17698 Human sec
573 45 93.8 85 7 ABJ72113 Abj72113 Human mem
574 45 93.8 85 7 ABR85599 Abr85599 Human sec
575 45 93.8 85 7 ABM77165 Abm77165 Human sec
576 45 93.8 85 7 ABO28344 Abo28344 Human sec
577 45 93.8 85 7 ABM23125 Abm23125 Human sec
578 45 93.8 85 7 ABM30445 Abm30445 Human sec
579 45 93.8 85 7 ABM21905 Abm21905 Human sec
580 45 93.8 85 7 ABM21600 Abm21600 Human sec
581 45 93.8 85 7 ABM15131 Abm15131 Human sec
582 45 93.8 85 7 ABO41206 Abo41206 Human sec
583 45 93.8 85 7 ABO36936 Abo36936 Human sec
584 45 93.8 85 7 ABO37546 Abo37546 Human sec
585 45 93.8 85 7 ABM75336 Abm75336 Human sec
586 45 93.8 85 7 ABM33616 Abm33616 Human sec
587 45 93.8 85 7 ABO46371 Abo46371 Human PRO
588 45 93.8 85 7 ADA82895 Ada82895 Human sec
589 45 93.8 85 7 ADB83602 Adb83602 Novel hum
590 45 93.8 85 7 ADB80708 Adb80708 Novel hum
591 45 93.8 85 7 ADB73249 Adb73249 Novel hum
592 45 93.8 85 7 ABM31970 Abm31970 Human sec
593 45 93.8 85 7 ADB78331 Adb78331 Novel hum
594 45 93.8 85 7 ABM31360 Abm31360 Human sec
595 45 93.8 85 7 ADB84979 Adb84979 Human PRO
596 45 93.8 85 7 ADB78085 Adb78085 Novel hum
597 45 93.8 85 7 ADB86203 Adb86203 Human sec
598 45 93.8 85 7 ABM32275 Abm32275 Human sec
599 45 93.8 85 7 ABM32580 Abm32580 Human sec
600 45 93.8 85 7 ADB87151 Adb87151 Human PRO
601 45 93.8 85 7 ADB84733 Adb84733 Human PRO
602 45 93.8 85 7 ABM31665 Abm31665 Human sec
603 45 93.8 85 7 ABM31055 Abm31055 Human sec
604 45 93.8 85 7 ADB83848 Adb83848 Novel hum
605 45 93.8 85 7 ADB73003 Adb73003 Novel hum
606 45 93.8 85 7 ADC36841 Adc36841 Human PRO
607 45 93.8 85 7 ADC21831 Adc21831 Human PRO
608 45 93.8 85 7 ADC49862 Adc49862 Novel hum

609 45 93.8 85 7 ADC49061 Adc49061 Novel hum
610 45 93.8 85 7 ADC49578 Adc49578 Novel hum
611 45 93.8 85 7 ADC47439 Adc47439 Novel hum
612 45 93.8 85 7 ADC47184 Adc47184 Novel hum
613 45 93.8 85 7 ADC78059 Adc78059 Novel hum
614 45 93.8 85 7 ADD06294 Add06294 Human sec
615 45 93.8 85 7 ADD05933 Add05933 Human sec
616 45 93.8 85 7 ADD10525 Add10525 Human sec
617 45 93.8 85 7 ADC77813 Adc77813 Novel hum
618 45 93.8 85 7 ADD11485 Add11485 Human sec
619 45 93.8 85 7 ADD50776 Add50776 Novel hum
620 45 93.8 85 7 ADD51022 Add51022 Novel hum
621 45 93.8 85 7 ADD37278 Add37278 Human sec
622 45 93.8 85 7 ADD50503 Add50503 Human PRO
623 45 93.8 85 7 ADD50257 Add50257 Human PRO
624 45 93.8 85 7 ADD51268 Add51268 Novel hum
625 45 93.8 85 7 ADG02928 Adg02928 Novel hum
626 45 93.8 85 7 ADG01635 Adg01635 Novel hum
627 45 93.8 85 7 ADF95810 Adf95810 Novel hum
628 45 93.8 85 7 ADG12625 Adg12625 Novel hum
629 45 93.8 85 7 ADH09285 Adh09285 Human PRO
630 45 93.8 85 7 ADG63737 Adg63737 Human sec
631 45 93.8 85 7 ADL33064 Adl33064 Novel hum
632 45 93.8 85 7 ADM30600 Adm30600 Novel hum
633 45 93.8 85 7 ADM77842 Adm77842 Human pro
634 45 93.8 85 8 ADC48815 Adc48815 Novel hum
635 45 93.8 85 8 ADE20986 Ade20986 Novel hum
636 45 93.8 85 8 ADE05830 Ade05830 Human PRO
637 45 93.8 85 8 ADD75059 Add75059 Human PRO
638 45 93.8 85 8 ADD75805 Add75805 Novel hum
639 45 93.8 85 8 ADD85037 Add85037 Novel hum
640 45 93.8 85 8 ADD86863 Add86863 Novel hum
641 45 93.8 85 8 ADE20740 Ade20740 Novel hum
642 45 93.8 85 8 ADE39037 Ade39037 Novel hum
643 45 93.8 85 8 ADE05584 Ade05584 Human PRO
644 45 93.8 85 8 ADD73569 Add73569 Human PRO
645 45 93.8 85 8 ADD78409 Add78409 Novel hum
646 45 93.8 85 8 ADE41486 Ade41486 Human sec
647 45 93.8 85 8 ADE74597 Ade74597 Human sec
648 45 93.8 85 8 ADE21232 Ade21232 Novel hum
649 45 93.8 85 8 ADD77347 Add77347 Novel hum
650 45 93.8 85 8 ADE20494 Ade20494 Novel hum
651 45 93.8 85 8 ADD75559 Add75559 Human PRO
652 45 93.8 85 8 ADD74075 Add74075 Human PRO
653 45 93.8 85 8 ADD74321 Add74321 Human PRO
654 45 93.8 85 8 ADD76051 Add76051 Novel hum
655 45 93.8 85 8 ADD85543 Add85543 Novel hum
656 45 93.8 85 8 ADE05092 Ade05092 Human PRO
657 45 93.8 85 8 ADD75305 Add75305 Human PRO
658 45 93.8 85 8 ADD76849 Add76849 Novel hum
659 45 93.8 85 8 ADD86617 Add86617 Novel hum
660 45 93.8 85 8 ADE41151 Ade41151 Human sec
661 45 93.8 85 8 ADD78085 Add78085 Novel hum
662 45 93.8 85 8 ADE75209 Ade75209 Human sec
663 45 93.8 85 8 ADD77593 Add77593 Novel hum
664 45 93.8 85 8 ADD77839 Add77839 Novel hum
665 45 93.8 85 8 ADD85297 Add85297 Novel hum
666 45 93.8 85 8 ADD73829 Add73829 Human PRO
667 45 93.8 85 8 ADD74567 Add74567 Human PRO
668 45 93.8 85 8 ADD77095 Add77095 Novel hum
669 45 93.8 85 8 ADD85789 Add85789 Novel hum
670 45 93.8 85 8 ADE05338 Ade05338 Human PRO
671 45 93.8 85 8 ADD74813 Add74813 Human PRO
672 45 93.8 85 8 ADG05625 Adg05625 Novel hum
673 45 93.8 85 8 ADG27179 Adg27179 Human PRO
674 45 93.8 85 8 ADF96422 Adf96422 Novel hum
675 45 93.8 85 8 ADG11242 Adg11242 Novel hum
676 45 93.8 85 8 ADG04693 Adg04693 Novel hum
677 45 93.8 85 8 ADG12021 Adg12021 Novel hum
678 45 93.8 85 8 ADG00853 Adg00853 Novel hum
679 45 93.8 85 8 ADF94578 Adf94578 Novel hum
680 45 93.8 85 8 ADG06674 Adg06674 Human PRO
681 45 93.8 85 8 ADG83109 Adg83109 Human PRO

682	45	93.8	85	8	ADH26390	Adh26390 Novel hum
683	45	93.8	85	8	ADH39018	Adh39018 Novel hum
684	45	93.8	85	8	ADG63586	Adg63586 Human sec
685	45	93.8	85	8	ADH33359	Adh33359 Human PRO
686	45	93.8	85	8	ADH43669	Adh43669 Human PRO
687	45	93.8	85	8	ADG34108	Adg34108 Novel hum
688	45	93.8	85	8	ADI33578	Adi33578 Human PRO
689	45	93.8	85	8	ADH69672	Adh69672 Human PRO
690	45	93.8	85	8	ADI29833	Adi29833 Novel hum
691	45	93.8	85	8	ADJ55098	Adj55098 Human PRO
692	45	93.8	85	8	ADM27230	Adm27230 Novel hum
693	45	93.8	85	8	ADK83014	Adk83014 Human PRO
694	45	93.8	85	8	ADJ64869	Adj64869 Human PRO
695	45	93.8	85	8	ADK66588	Adk66588 Human PRO
696	45	93.8	85	8	ADM31765	Adm31765 Novel hum
697	45	93.8	85	8	ADM36812	Adm36812 Novel hum
698	45	93.8	85	8	ADM40617	Adm40617 Novel hum
699	45	93.8	85	8	ADN38225	Adn38225 Novel hum
700	45	93.8	85	8	ADP19523	Adp19523 Human sec
701	45	93.8	85	8	ADP19346	Adp19346 Human sec
702	45	93.8	85	8	ADP19345	Adp19345 Human sec
703	45	93.8	85	8	ADU73028	Adu73028 cDNA 48-1
704	45	93.8	85	9	ADZ89302	Adz89302 Secreted
705	45	93.8	86	2	AAY36257	Aay36257 Human sec
706	45	93.8	86	6	ADAL1687	Adal1687 Human nov
707	45	93.8	122	6	ADA12081	Adal2081 Human nov
708	44	91.7	12	2	AAR93355	Aar93355 FYN prote
709	44	91.7	112	8	ABM80911	Abm80911 Tumour-as
710	44	91.7	147	8	ADR09417	Adr09417 Human pro
711	43	89.6	12	2	AAR93347	Aar93347 SRC prote
712	43	89.6	209	7	ADI62971	Adi62971 Human apo
713	43	89.6	209	7	ADI63050	Adi63050 Human apo
714	43	89.6	318	4	ABG19764	Abg19764 Novel hum
715	43	89.6	462	3	AAB07572	Aab07572 Protein e
716	43	89.6	514	8	ADN05033	Adn05033 Antipsori
717	43	89.6	514	8	ADP25283	Adp25283 PRO polyp
718	43	89.6	1268	4	ABB67095	Abb67095 Drosophil
719	42	87.5	19	2	AAW05473	Aaw05473 SH3-bindi
720	42	87.5	127	5	AAE29154	Aae29154 Human non
721	42	87.5	158	7	ABO76210	AbO76210 Pseudomon
722	42	87.5	170	7	ABO76223	AbO76223 Pseudomon
723	42	87.5	173	8	ADY22951	Ady22951 Plant ful
724	42	87.5	210	8	ADX87621	Adx87621 Plant ful
725	42	87.5	213	8	ADX88446	Adx88446 Plant ful
726	42	87.5	258	2	AAV04371	Aay04371 Human Fas
727	42	87.5	258	7	ADG14424	Adg14424 Human Fas
728	42	87.5	261	2	AAY28597	Aay28597 Fas ligan
729	42	87.5	265	2	AAW48954	Aaw48954 Non-cleav
730	42	87.5	267	8	ADJ63977	Adj63977 Human Fas
731	42	87.5	267	8	ADL71817	Adl71817 Human Fas
732	42	87.5	267	8	ADN07588	Adn07588 Human Fas
733	42	87.5	268	2	AAW48953	Aaw48953 Non-cleav
734	42	87.5	271	2	AAY28596	Aay28596 Fas ligan
735	42	87.5	277	2	AAV04372	Aay04372 Human Fas
736	42	87.5	277	2	AAY28595	Aay28595 Fas ligan
737	42	87.5	278	6	ABR82203	Abr82203 Chimeric
738	42	87.5	280	8	ADJ36211	Adj36211 Self-coal
739	42	87.5	281	2	AAR77281	Aar77281 Human Fas
740	42	87.5	281	2	AAR79097	Aar79097 Human Fas
741	42	87.5	281	2	AAR88356	Aar88356 Human Fas
742	42	87.5	281	2	AAR98104	Aar98104 Human Fas
743	42	87.5	281	2	AAW27143	Aaw27143 Human Fas
744	42	87.5	281	2	AAW75959	Aaw75959 Human Fas
745	42	87.5	281	2	AAW49105	Aaw49105 Fas Ligan
746	42	87.5	281	2	AAW98071	Aaw98071 Human Fas
747	42	87.5	281	2	AAV04373	Aay04373 Human Fas
748	42	87.5	281	2	AAW95041	Aaw95041 Human Fas
749	42	87.5	281	2	AAY28594	Aay28594 Wild type
750	42	87.5	281	3	AAY87580	Aay87580 Human Fas
751	42	87.5	281	3	AAY87577	Aay87577 Human Fas
752	42	87.5	281	3	AAY87569	Aay87569 Human Fas
753	42	87.5	281	3	AAV87575	Aay87575 Human Fas
754	42	87.5	281	3	AAY87576	Aay87576 Human Fas

755	42	87.5	281	3	AAV87579	Aay87579 Human Fas
756	42	87.5	281	3	AAV87574	Aay87574 Human Fas
757	42	87.5	281	3	AAV87582	Aay87582 Human Fas
758	42	87.5	281	3	AAV87578	Aay87578 Human Fas
759	42	87.5	281	3	AAV87581	Aay87581 Human Fas
760	42	87.5	281	3	AAB19342	Aab19342 Amino aci
761	42	87.5	281	4	AAU04556	Aau04556 Human Fas
762	42	87.5	281	5	ABB80000	Abb80000 Human Fas
763	42	87.5	281	5	ABG96462	Abg96462 Human Fas
764	42	87.5	281	5	ABB05000	Abb05000 Human Fas
765	42	87.5	281	5	ABB81652	Abb81652 Human Fas
766	42	87.5	281	5	AAE29153	Aae29153 Wild-type
767	42	87.5	281	6	ABR42309	Abr42309 Human Fas
768	42	87.5	281	6	ABP60549	Abp60549 Human tum
769	42	87.5	281	7	ADC35194	Adc35194 Human TNF
770	42	87.5	281	7	ADF44995	Adf44995 Human Fas
771	42	87.5	281	7	ABW02272	Abw02272 Human Fas
772	42	87.5	281	7	ADG14422	Adg14422 Human Fas
773	42	87.5	281	7	ADJ74024	Adj74024 Human Fas
774	42	87.5	281	7	ADJ92614	Adj92614 Human Fas
775	42	87.5	281	7	ADL17694	Adl17694 Human Fas
776	42	87.5	281	8	ADF90414	Adf90414 Human Fas
777	42	87.5	281	8	ADJ56767	Adj56767 Human Fas
778	42	87.5	281	8	ADL27664	Adl27664 Human Fas
779	42	87.5	281	8	ADL23577	Adl23577 Human Fas
780	42	87.5	281	8	ADK19674	Adk19674 Chicken a
781	42	87.5	281	8	ADK19673	Adk19673 Human Fas
782	42	87.5	281	8	ADM53436	Adm53436 Human Fas
783	42	87.5	281	8	ADP12464	Adp12464 Protein e
784	42	87.5	281	8	ADO19786	Ado19786 Human PRO
785	42	87.5	281	8	ADO19800	Ado19800 Human PRO
786	42	87.5	281	8	ADP26984	Adp26984 Human Fas
787	42	87.5	281	8	ADS87994	Ads87994 Tumour tr
788	42	87.5	281	8	ADP56004	Adp56004 Human PRO
789	42	87.5	281	8	ADR44815	Adr44815 FAS ligan
790	42	87.5	281	8	ADR47163	Adr47163 Human Fas
791	42	87.5	281	8	ADT78394	Adt78394 Human Fas
792	42	87.5	281	8	ADU22865	Adu22865 Human apo
793	42	87.5	281	9	ADY19512	Ady19512 PRO polyp
794	42	87.5	281	9	ADY16498	Ady16498 PRO polyp
795	42	87.5	281	9	ADZ14447	Adz14447 Human Fas
796	42	87.5	329	8	ADJ36255	Adj36255 Self-coal
797	42	87.5	347	8	ADY05457	Ady05457 Plant ful
798	42	87.5	354	7	ABM89848	Abm89848 Rice abio
799	42	87.5	434	4	AAM42098	Aam42098 Human pol
800	42	87.5	853	4	ABG12054	Abg12054 Novel hum
801	41	85.4	7	2	AAW11128	Aaw11128 Src SH3 d
802	41	85.4	7	2	AAW17010	Aaw17010 SRC SH3 d
803	41	85.4	7	2	AAW25486	Aaw25486 SH3 domai
804	41	85.4	7	2	AAW79781	Aaw79781 Proline-r
805	41	85.4	7	3	AAB17267	Aab17267 Src antag
806	41	85.4	7	3	AAB17226	Aab17226 SH3 antag
807	41	85.4	7	3	AAY69979	Aay69979 Src SH3 r
808	41	85.4	7	4	AAB50762	Aab50762 Human cAM
809	41	85.4	7	5	ABB73219	Abb73219 Src homol
810	41	85.4	7	5	ABB73345	Abb73345 Exemplary
811	41	85.4	7	7	ADJ73499	Adj73499 Exemplary
812	41	85.4	7	7	ADJ73373	Adj73373 SH3 antag
813	41	85.4	7	8	ADJ53007	Adj53007 CH1 delet
814	41	85.4	7	8	ADJ53133	Adj53133 CH1 delet
815	41	85.4	7	8	ADJ51968	Adj51968 CH1 delet
816	41	85.4	7	8	ADJ52094	Adj52094 CH1 delet
817	41	85.4	10	2	AAR93545	Aar93545 Random 10
818	41	85.4	10	2	AAR93544	Aar93544 Random 10
819	41	85.4	10	9	AEB07369	Aeb07369 Signal tr
820	41	85.4	12	2	AAR93378	Aar93378 Grb-2 pro
821	41	85.4	12	2	AAR93380	Aar93380 Grb-2 pro
822	41	85.4	12	2	AAR93352	Aar93352 FYN prote
823	41	85.4	12	2	AAR93353	Aar93353 FYN prote
824	41	85.4	12	2	AAR93349	Aar93349 FYN prote
825	41	85.4	12	2	AAR93379	Aar93379 Grb-2 pro
826	41	85.4	12	2	AAR93364	Aar93364 LYN prote
827	41	85.4	12	2	AAR93365	Aar93365 LYN prote

828 41 85.4 12 2 AAR933360 Aar933360 LYN prote 901 41 85.4 230 8 ADY06849 Ady06849 Plant ful
829 41 85.4 12 2 AAR933363 Aar933363 LYN prote 902 41 85.4 278 5 ABB78672 Abb78672 C. elegan
830 41 85.4 12 2 AAR933343 Aar933343 SRC prote 903 41 85.4 288 8 ADX79901 Adx79901 Plant ful
831 41 85.4 12 2 AAR933348 Aar933348 SRC prote 904 41 85.4 292 4 AAB61267 Aab61267 Mature mo
832 41 85.4 12 3 AAB17253 Aab17253 SH3 antag 905 41 85.4 292 6 ABU11233 Abu11233 Mouse TAN
833 41 85.4 12 4 AAB50782 Aab50782 Human cAM 906 41 85.4 296 8 ADI45297 Adi45297 Rice isop
834 41 85.4 12 5 ABB73246 Abb73246 Src homol 907 41 85.4 303 5 ABB78673 Abb78673 C. elegan
835 41 85.4 12 6 ADA08218 Ada08218 Human Src 908 41 85.4 313 4 AAB61265 Aab61265 Mouse TAN
836 41 85.4 12 7 ADJ73400 Adj73400 SH3 antag 909 41 85.4 313 4 AAB61278 Aab61278 Mouse TAN
837 41 85.4 12 8 ADJ53034 Adj53034 CH1 delet 910 41 85.4 313 4 AAB61280 Aab61280 Mouse TAN
838 41 85.4 12 8 ADJ51995 Adj51995 CH1 delet 911 41 85.4 313 4 AAB61277 Aab61277 Mouse TAN
839 41 85.4 13 2 AAW05483 Aaw05483 SH3-bindi 912 41 85.4 313 4 AAB61279 Aab61279 Mouse TAN
840 41 85.4 14 2 AAW11099 Aaw11099 Src SH3 d 913 41 85.4 313 6 ABU11246 Abu11246 Glycoprot
841 41 85.4 14 2 AAW25517 Aaw25517 SH3 synth 914 41 85.4 313 6 ABU11245 Abu11245 Glycoprot
842 41 85.4 14 3 AAB08530 Aab08530 Amino aci 915 41 85.4 313 6 ABU11243 Abu11243 Glycoprot
843 41 85.4 14 4 AAB36691 Aab36691 Pro5 pept 916 41 85.4 313 6 ABU11231 Abu11231 Mouse TAN
844 41 85.4 14 5 AAG66171 Aag66171 c-Src pep 917 41 85.4 313 6 ABU11244 Abu11244 Glycoprot
845 41 85.4 15 2 AAR933470 Aar933470 GST-SRC p 918 41 85.4 319 4 AAB61257 Aab61257 Mature hu
846 41 85.4 15 2 AAR933463 Aar933463 GST-PI3K 919 41 85.4 319 4 AAB49403 Aab49403 Human gly
847 41 85.4 15 2 AAR933465 Aar933465 GST-PI3K 920 41 85.4 319 6 ABU11223 Abu11223 Human TAN
848 41 85.4 15 2 AAR933464 Aar933464 GST-PI3K 921 41 85.4 321 8 ADP08399 Adp08399 Human gly
849 41 85.4 15 2 AAR933462 Aar933462 GST-PI3K 922 41 85.4 321 8 ADQ39541 Adq39541 Human myo
850 41 85.4 15 2 AAR933466 Aar933466 GST-PI3K 923 41 85.4 325 5 ABB78671 Abb78671 C. elegan
851 41 85.4 15 2 AAW05069 Aaw05069 Proline-r 924 41 85.4 339 4 AAB61273 Aab61273 Human TAN
852 41 85.4 15 2 AAW05066 Aaw05066 Proline-r 925 41 85.4 339 4 AAB61255 Aab61255 Human TAN
853 41 85.4 15 2 AAW05070 Aaw05070 Proline-r 926 41 85.4 339 4 AAB61274 Aab61274 Human TAN
854 41 85.4 15 2 AAW38925 Aaw38925 Peptide r 927 41 85.4 339 4 AAB61275 Aab61275 Human TAN
855 41 85.4 15 2 AAW38966 Aaw38966 Peptide r 928 41 85.4 339 4 AAB61276 Aab61276 Human TAN
856 41 85.4 15 2 AAW38958 Aaw38958 Peptide r 929 41 85.4 339 4 AAY72790 Aay72790 Human pla
857 41 85.4 15 3 AAB20793 Aab20793 Src, Fyn, 930 41 85.4 339 4 AAB31668 Aab31668 Amino aci
858 41 85.4 16 2 AAW25448 Aaw25448 Grb2 N-te 931 41 85.4 339 5 ABB78674 Abb78674 C. elegan
859 41 85.4 21 2 AAW05452 Aaw05452 SH3-bindi 932 41 85.4 339 6 ADA84105 Ada84105 Human GP6
860 41 85.4 22 2 AAW37660 Aaw37660 PPPPY mot 933 41 85.4 339 6 ABU11221 Abu11221 Human TAN
861 41 85.4 22 7 ADB49286 Adb49286 Novel WW 934 41 85.4 339 6 ABU11239 Abu11239 Glycoprot
862 41 85.4 25 9 AEB07370 Aeb07370 Signal tr 935 41 85.4 339 6 ABU11240 Abu11240 Glycoprot
863 41 85.4 27 4 AAB61270 Aab61270 Mouse TAN 936 41 85.4 339 6 ABU11241 Abu11241 Glycoprot
864 41 85.4 27 6 ABU11236 Abu11236 Mouse TAN 937 41 85.4 339 6 ABU11242 Abu11242 Glycoprot
865 41 85.4 30 9 AEB07371 Aeb07371 Signal tr 938 41 85.4 339 8 ADH22420 Adh22420 Human pro
866 41 85.4 40 4 AAM18998 Aam18998 Peptide # 939 41 85.4 339 8 ADP08400 Adp08400 Human gly
867 41 85.4 40 4 ABB38188 Abb38188 Peptide # 940 41 85.4 386 5 ABB78660 Abb78660 C. elegan
868 41 85.4 40 4 AAM31617 Aam31617 Peptide # 941 41 85.4 386 5 ABB78670 Abb78670 C. elegan
869 41 85.4 40 4 ABB23379 Aab23379 Protein # 942 41 85.4 386 5 ABB78668 Abb78668 C. elegan
870 41 85.4 40 4 AAM71338 Aam71338 Human bon 943 41 85.4 386 5 ABB78675 Abb78675 C. elegan
871 41 85.4 40 4 AAM58822 Aam58822 Human bra 944 41 85.4 386 5 ABB78665 Abb78665 C. elegan
872 41 85.4 40 4 ABG53046 Abg53046 Human liv 945 41 85.4 386 5 ABB78648 Abb78648 Caenorhab
873 41 85.4 40 5 ABG41141 Abg41141 Human pep 946 41 85.4 386 5 ABB78667 Abb78667 C. elegan
874 41 85.4 51 4 AAB61262 Aab61262 Human TAN 947 41 85.4 386 5 ABB78664 Abb78664 C. elegan
875 41 85.4 51 6 ABU11228 Abu11228 Human TAN 948 41 85.4 386 5 ABB78661 Abb78661 C. elegan
876 41 85.4 73 5 ABB09414 Abb09414 Ovarian c 949 41 85.4 398 9 AEB20718 Aeb20718 TRP domai
877 41 85.4 73 5 AAO18093 Aao18093 Human ova 950 41 85.4 398 9 AEB07376 Aeb07376 Signal tr
878 41 85.4 73 7 ADM10871 Adm10871 O573S 411 951 41 85.4 401 4 AAM79810 Aam79810 Human pro
879 41 85.4 73 8 ADJ11201 Adj11201 Human_ova 952 41 85.4 432 8 ADX87392 Adx87392 Plant ful
880 41 85.4 73 8 ADM43462 Adm43462 Human ova 953 41 85.4 459 7 ADE63349 Ade63349 Rat Prote
881 41 85.4 78 5 ABB78662 Abb78662 C. elegan 954 41 85.4 461 8 ADO60064 Ado60064 CRH signa
882 41 85.4 83 3 AAB40669 Aab40669 Human ORF 955 41 85.4 462 3 AAB18668 Aab18668 A human r
883 41 85.4 83 5 ABP06901 Abp06901 Human ORF 956 41 85.4 462 4 AAM78826 Aam78826 Human pro
884 41 85.4 89 8 ADX66827 Adx66827 Plant ful 957 41 85.4 462 5 ABG70175 Abg70175 Human pre
885 41 85.4 108 4 AAB36682 Aab36682 Mammalian 958 41 85.4 462 6 ABUS7649 Abu57649 Different
886 41 85.4 112 5 ABB78663 Abb78663 C. elegan 959 41 85.4 462 7 ADE63351 Ade63351 Human pro
887 41 85.4 127 4 AAO01550 Aao01550 Human pol 960 41 85.4 462 8 ADL12698 Adl12698 Human ste
888 41 85.4 146 5 ABB78666 Abb78666 C. elegan 961 41 85.4 480 4 AAB36684 Aab36684 Mammalian
889 41 85.4 152 8 ADY22975 Ady22975 Plant ful 962 41 85.4 495 6 ABR92076 Abr92076 Human cer
890 41 85.4 156 2 AAR13587 Aar13587 PLRV vira 963 41 85.4 495 7 ADE34349 Ade34349 Human hyp
891 41 85.4 161 4 AAB36683 Aab36683 Mammalian 964 41 85.4 495 8 ADP24979 Adp24979 PRO polyp
892 41 85.4 162 5 ABB78669 Abb78669 C. elegan 965 41 85.4 495 8 ADU06363 Adu06363 Novel bro
893 41 85.4 176 3 AAY50994 Aay50994 Potato pr 966 41 85.4 513 6 ABP96232 Abp96232 Human nuc
894 41 85.4 183 8 ADX90517 Adx90517 Plant ful 967 41 85.4 535 4 ABB66593 Abb66593 Drosophil
895 41 85.4 186 8 ADK48382 Adk48382 Streptoco 968 41 85.4 548 4 ABB59289 Abb59289 Drosophil
896 41 85.4 186 8 ADX90922 Adx90922 Plant ful 969 41 85.4 562 5 ABP69088 Abp69088 Human pol
897 41 85.4 203 4 ABG15907 Abg15907 Novel hum 970 41 85.4 710 4 AAM43571 Aam43571 Human pol
898 41 85.4 206 1 ABG13470 Abg13470 Novel hum 971 41 85.4 710 8 ADM24592 Adm24592 Human PRO
899 41 85.4 219 8 ADX79537 Adx79537 Plant ful 972 41 85.4 717 4 AAU18383 Aau18383 Human end
900 41 85.4 225 4 AAB36690 Aab36690 Mammalian 973 41 85.4 760 4 AAM93625 Aam93625 Human pol

901 41 85.4 230 8 ADY06849 Ady06849 Plant ful
902 41 85.4 278 5 ABB78672 Abb78672 C. elegan
903 41 85.4 288 8 ADX79901 Adx79901 Plant ful
904 41 85.4 292 4 AAB61267 Aab61267 Mature mo
905 41 85.4 292 6 ABU11233 Abu11233 Mouse TAN
906 41 85.4 296 8 ADI45297 Adi45297 Rice isop
907 41 85.4 303 5 ABB78673 Abb78673 C. elegan
908 41 85.4 313 4 AAB61265 Aab61265 Mouse TAN
909 41 85.4 313 4 AAB61278 Aab61278 Mouse TAN
910 41 85.4 313 4 AAB61280 Aab61280 Mouse TAN
911 41 85.4 313 4 AAB61277 Aab61277 Mouse TAN
912 41 85.4 313 4 AAB61279 Aab61279 Mouse TAN
913 41 85.4 313 6 ABU11246 Abu11246 Glycoprot
914 41 85.4 313 6 ABU11245 Abu11245 Glycoprot
915 41 85.4 313 6 ABU11243 Abu11243 Glycoprot
916 41 85.4 313 6 ABU11231 Abu11231 Mouse TAN
917 41 85.4 313 6 ABU11244 Abu11244 Glycoprot
918 41 85.4 319 4 AAB61257 Aab61257 Mature hu
919 41 85.4 319 4 AAB49403 Aab49403 Human gly
920 41 85.4 319 6 ABU11223 Abu11223 Human TAN
921 41 85.4 321 8 ADP08399 Adp08399 Human gly
922 41 85.4 321 8 ADQ39541 Adq39541 Human myo
923 41 85.4 325 5 ABB78671 Abb78671 C. elegan
924 41 85.4 339 4 AAB61273 Aab61273 Human TAN
925 41 85.4 339 4 AAB61255 Aab61255 Human TAN
926 41 85.4 339 4 AAB61274 Aab61274 Human TAN
927 41 85.4 339 4 AAB61275 Aab61275 Human TAN
928 41 85.4 339 4 AAB61276 Aab61276 Human TAN
929 41 85.4 339 4 AAY72790 Aay72790 Human pla
930 41 85.4 339 4 AAB31668 Aab31668 Amino aci
931 41 85.4 339 5 ABB78674 Abb78674 C. elegan
932 41 85.4 339 6 ADA84105 Ada84105 Human GP6
933 41 85.4 339 6 ABU11221 Abu11221 Human TAN
934 41 85.4 339 6 ABU11239 Abu11239 Glycoprot
935 41 85.4 339 6 ABU11240 Abu11240 Glycoprot
936 41 85.4 339 6 ABU11241 Abu11241 Glycoprot
937 41 85.4 339 6 ABU11242 Abu11242 Glycoprot
938 41 85.4 339 8 ADH22420 Adh22420 Human pro
939 41 85.4 339 8 ADP08400 Adp08400 Human gly
940 41 85.4 386 5 ABB78660 Abb78660 C. elegan
941 41 85.4 386 5 ABB78670 Abb78670 C. elegan
942 41 85.4 386 5 ABB78668 Abb78668 C. elegan
943 41 85.4 386 5 ABB78675 Abb78675 C. elegan
944 41 85.4 386 5 ABB78665 Abb78665 C. elegan
945 41 85.4 386 5 ABB78648 Abb78648 Caenorhab
946 41 85.4 386 5 ABB78667 Abb78667 C. elegan
947 41 85.4 386 5 ABB78664 Abb78664 C. elegan
948 41 85.4 386 5 ABB78661 Abb78661 C. elegan
949 41 85.4 398 9 AEB20718 Aeb20718 TRP domai
950 41 85.4 398 9 AEB07376 Aeb07376 Signal tr
951 41 85.4 401 4 AAM79810 Aam79810 Human pro
952 41 85.4 432 8 ADX87392 Adx87392 Plant ful
953 41 85.4 459 7 ADE63349 Ade63349 Rat Prote
954 41 85.4 461 8 ADO60064 Ado60064 CRH signa
955 41 85.4 462 3 AAB18668 Aab18668 A human r
956 41 85.4 462 4 AAM78826 Aam78826 Human pro
957 41 85.4 462 5 ABG70175 Abg70175 Human pre
958 41 85.4 462 6 ABUS7649 Abu57649 Different
959 41 85.4 462 7 ADE63351 Ade63351 Human pro
960 41 85.4 462 8 ADL12698 Adl12698 Human ste
961 41 85.4 480 4 AAB36684 Aab36684 Mammalian
962 41 85.4 495 6 ABR92076 Abr92076 Human cer
963 41 85.4 495 7 ADE34349 Ade34349 Human hyp
964 41 85.4 495 8 ADP24979 Adp24979 PRO polyp
965 41 85.4 495 8 ADU06363 Adu06363 Novel bro
966 41 85.4 513 6 ABP96232 Abp96232 Human nuc
967 41 85.4 535 4 ABB66593 Abb66593 Drosophil
968 41 85.4 548 4 ABB59289 Abb59289 Drosophil
969 41 85.4 562 5 ABP69088 Abp69088 Human pol
970 41 85.4 710 4 AAM43571 Aam43571 Human pol
971 41 85.4 710 8 ADM24592 Adm24592 Human PRO
972 41 85.4 717 4 AAU18383 Aau18383 Human end
973 41 85.4 760 4 AAM93625 Aam93625 Human pol

974 41 85.4 760 8 ADL31429 Adl31429 Human pro
975 41 85.4 782 8 ADT56202 Adt56202 Plant pol
976 41 85.4 847 5 ABG34076 Abg34076 Human PRO
977 41 85.4 847 6 ADA01362 Ada01362 Human PRO
978 41 85.4 847 6 ADA43791 Ada43791 Human sec
979 41 85.4 847 6 ADA43559 Ada43559 Human sec
980 41 85.4 847 6 ADA01234 Ada01234 Human PRO
981 41 85.4 847 7 ADA01118 Ada01118 Human sec
982 41 85.4 847 7 ADA43675 Ada43675 Human sec
983 41 85.4 847 7 ADA06937 Ada06937 Human PRO
984 41 85.4 847 7 ADA08425 Ada08425 Novel hum
985 41 85.4 847 7 ADB99718 Adb99718 Human PRO
986 41 85.4 847 7 ADB87001 Adb87001 Human PRO
987 41 85.4 847 7 ADB66156 Adb66156 Human sec
988 41 85.4 847 7 ADB99834 Adb99834 Human PRO
989 41 85.4 847 7 ADB99489 Adb99489 Novel hum
990 41 85.4 847 7 ADB66040 Adb66040 Human sec
991 41 85.4 847 7 ADC23438 Adc23438 Human tra
992 41 85.4 847 7 ADC26131 Adc26131 Human PRO
993 41 85.4 847 7 ADE04958 Ade04958 Human PRO
994 41 85.4 847 7 ADE11264 Ade11264 Human PRO
995 41 85.4 847 7 ADD88195 Add88195 Human PRO
996 41 85.4 847 7 ADD95490 Add95490 Human sec
997 41 85.4 847 7 ADE06420 Ade06420 Human PRO
998 41 85.4 847 7 ADE38195 Ade38195 Human PRO
999 41 85.4 847 7 ADD88311 Add88311 Human PRO
1000 41 85.4 847 7 ADD90892 Add90892 Human sec

ALIGNMENTS

RESULT 1
AAR933346
ID AAR933346 standard; peptide; 12 AA.
XX
AC AAR933346;
XX
DT 24-APR-1996 (first entry)
XX
DE SRC/LYN protein tyrosine kinase derived peptide.
XX
KW SH3 ligand; SH3 binding agent; biased phage library;
KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
KW p67; complex; chronic myelogenous leukaemia; cancer.
XX
OS Synthetic.
XX
PN WO9524419-A1.
XX
PD 14-SEP-1995.
XX
PF 13-MAR-1995; 95WO-US003208.
XX
PR 11-MAR-1994; 94US-00209835.
PR 06-JAN-1995; 95US-00369832.
XX
PA (ARIA-) ARIAD PHARM INC.
XX
PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
XX
DR WPI; 1995-328231/42.
XX
PT Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.
XX
PS Disclosure; Fig 2; 74pp; English.
XX
CC The sequences given in AAR933343-68 represent peptides which are SH3
CC ligands/SH3 binding agents. They represent a biased phage library which
CC comprises six random amino acids flanking the hexapeptide RSLRPL- which

CC was identified as a recognition sequence for the src SH3 domain. These
CC sequences were identified using the method of the invention. The method
CC comprises contacting the SH3 domain with a mixture of peptides under
CC conditions permitting a ligand to bind to an SH3 domain to form a
CC complex. Any unbound peptides are removed and the complexed peptide
CC ligands are dissociated from the complexes. The selected peptides are
CC enriched by re-contacting them with the SH3 domain and then candidates
CC which bind to the SH3 domain are detected. The isolated SH3 binding
CC peptides may be used in the diagnosis, prevention and treatment of
CC conditions or diseases resulting from cellular processes mediated by an
CC SH3-based interaction. Such diseases include Paget's disease. Other
CC conditions treatable with these peptides include restenosis, rheumatoid
CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
CC p47 and p67 complex is implicated, etc
XX
SQ Sequence 12 AA;
Query Match 100.0%; Score 48; DB 2; Length 12;
Best Local Similarity 88.9%; Pred.No. 6.3;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 RPLPPLPXP 12
|||
Db 4 RPLPPLPSP 12
RESULT 2
AAR933356
ID AAR933356 standard; peptide; 12 AA.
XX
AC AAR933356;
XX
DT 24-APR-1996 (first entry)
XX
DE FYN/LYN protein tyrosine kinase derived peptide.
XX
KW SH3 ligand; SH3 binding agent; biased phage library;
KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
KW p67; complex; chronic myelogenous leukaemia; cancer.
XX
OS Synthetic.
XX
PN WO9524419-A1.
XX
PD 14-SEP-1995.
XX
PF 13-MAR-1995; 95WO-US003208.
XX
PR 11-MAR-1994; 94US-00209835.
PR 06-JAN-1995; 95US-00369832.
XX
PA (ARIA-) ARIAD PHARM INC.
XX
PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
XX
DR WPI; 1995-328231/42.
XX
PT Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.
XX
PS Disclosure; Fig 2; 74pp; English.
XX
CC The sequences given in AAR933343-68 represent peptides which are SH3
CC ligands/SH3 binding agents. They represent a biased phage library which
CC comprises six random amino acids flanking the hexapeptide RSLRPL- which
CC was identified as a recognition sequence for the src SH3 domain. These
CC sequences were identified using the method of the invention. The method
CC comprises contacting the SH3 domain with a mixture of peptides under
CC conditions permitting a ligand to bind to an SH3 domain to form a
CC complex. Any unbound peptides are removed and the complexed peptide
CC ligands are dissociated from the complexes. The selected peptides are

CC enriched by re-contacting them with the SH3 domain and then candidates
CC which bind to the SH3 domain are detected. The isolated SH3 binding
CC peptides may be used in the diagnosis, prevention and treatment of
CC conditions or diseases resulting from cellular processes mediated by an
CC SH3-based interaction. Such diseases include Paget's disease. Other
CC conditions treatable with these peptides include restenosis, rheumatoid
CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
CC p47 and p67 complex is implicated, etc
XX
SQ Sequence 12 AA;
Query Match 100.0%; Score 48; DB 2; Length 12;
Best Local Similarity 88.9%; Pred. No. 6.3;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 RPLPPLXP 12
Db 4 RPLPPLPAP 12
RESULT 3
AAR93471
ID AAR93471 standard; peptide; 15 AA.
XX
AC AAR93471;
XX 09-MAY-1996 (first entry)
DT
XX GST-SRC protein tyrosine kinase derived peptide #5.
DE
XX SH3 ligand; SH3 binding agent; biased phage library;
KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
KW p67; complex; chronic myelogenous leukaemia; cancer.
XX
OS Synthetic.
XX WO9524419-A1.
PN
XX 14-SEP-1995.
PD
XX
XX 13-MAR-1995; 95WO-US003208.
PF
XX
XX 11-MAR-1994; 94US-00209835.
PR
XX 06-JAN-1995; 95US-00369832.
PR
XX (ARIA-) ARIAD PHARM INC.
PA
XX Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
PI
XX WPI; 1995-328231/42.
DR
XX Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.
PT
XX Disclosure; Fig 5; 74pp; English.
PS
XX The sequences given in AAR93457-71 represent peptides which are SH3
CC ligands/SH3 binding agents. They represent a biased phage library which
CC comprises five random amino acids flanking the decapeptide -RSLRPLPLP or
CC derivatives of this, which was identified as a recognition sequence for
CC the src SH3 domain. These sequences were identified using the method of
CC the invention. The method comprises contacting the SH3 domain with a
CC mixture of peptides under conditions permitting a ligand to bind to an
CC SH3 domain to form a complex. Any unbound peptides are removed and the
CC complexed peptide ligands are dissociated from the complexes. The
CC selected peptides are enriched by re-contacting them with the SH3 domain
CC and then candidates which bind to the SH3 domain are detected. The
CC isolated SH3 binding peptides may be used in the diagnosis, prevention
CC and treatment of conditions or diseases resulting from cellular processes
CC mediated by an SH3-based interaction. Such diseases include Paget's
CC disease. Other conditions treatable with these peptides include

CC restenosis, rheumatoid arthritis, gout and other problems in which an SH3
CC of neutrophil oxidase p47 and p67 complex is implicated, etc
XX
SQ Sequence 15 AA;
Query Match 100.0%; Score 48; DB 2; Length 15;
Best Local Similarity 88.9%; Pred. No. 7.7;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 RPLPPLXP 12
Db 4 RPLPPLPTP 12
RESULT 4
AAB21130
ID AAB21130 standard; peptide; 11 AA.
XX
AC AAB21130;
XX
DT 19-JAN-2001 (first entry)
XX Src homology 3 domain binding peptide #7.
DE
XX Src homology domain 3; SH3; protein-protein interaction; cancer;
KW signal transduction inhibition; immune suppression-associated disease.
KW
XX Synthetic.
OS
XX WO200047607-A1.
PN
XX 17-AUG-2000.
PD
XX 12-FEB-2000; 2000WO-KR000107.
PF
XX 12-FEB-1999; 99AU-00008643.
PR
XX 02-JUN-1999; 99KR-00020282.
PR
XX (YOON/) YOON J H.
PA (HANY/) HAN Y T.
PA
XX Yoon JH, Han YT, Lee KY;
PI
XX WPI; 2000-533010/48.
DR
XX Synthetic peptides useful for treating cancers and immunosuppressive
PT disorders by disrupting interactions of the SH (Src homology) 2 and SH3
PT motifs of Src family kinase proteins.
PT
XX Claim 3; Page 34; 40pp; English.
PS
XX The present sequence is a synthetic peptide which has a high affinity for
CC the src homology 3 (SH3) domain of protein kinases. Protein kinases are
CC involved in signal transduction pathways, and this peptide can be used to
CC inhibit these, by disrupting protein-protein interactions, in the
CC treatment of cancer, particularly hepatocellular carcinoma, cervical
CC cancer, colon adenocarcinoma and fibrosarcoma, and immune suppression-
CC associated diseases
XX
SQ Sequence 11 AA;
Query Match 97.9%; Score 47; DB 3; Length 11;
Best Local Similarity 88.9%; Pred. No. 7.9;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 RPLPPLXP 12
Db 2 RPLPPLPLP 10
RESULT 5
AAB21125
ID AAB21125 standard; peptide; 11 AA.

XX AAB211125;
AC 19-JAN-2001 (first entry)
XX
XX Src homology 3 domain binding peptide #2.
DE
XX Src homology domain 3; SH3; protein-protein interaction; cancer;
KW signal transduction inhibition; immune suppression-associated disease.
KW
XX Synthetic.
OS
XX WO200047607-A1.
PN
XX 17-AUG-2000.
PD
XX
XX 12-FEB-2000; 2000WO-KR000107.
PF
XX
XX 12-FEB-1999; 99AU-00008643.
PR
XX 02-JUN-1999; 99KR-00020282.
PR
XX (YOON/) YOON J H.
PA (HANY/) HAN Y T.
PA
XX Yoon JH, Han YT, Lee KY;
PI
XX WPI; 2000-533010/48.
DR
XX Synthetic peptides useful for treating cancers and immunosuppressive
PT disorders by disrupting interactions of the SH (Src homology) 2 and SH3
PT motifs of Src family kinase proteins.
PT
XX Claim 3; Page 32; 40pp; English.
PS
XX The present sequence is a synthetic peptide which has a high affinity for
CC the src homology 3 (SH3) domain of protein kinases. Protein kinases are
CC involved in signal transduction pathways, and this peptide can be used to
CC inhibit these, by disrupting protein-protein interactions, in the
CC treatment of cancer, particularly hepatocellular carcinoma, cervical
CC cancer, colon adenocarcinoma and fibrosarcoma, and immune suppression-
CC associated diseases
XX
SQ Sequence 11 AA;
Query Match 97.9%; Score 47; DB 3; Length 11;
Best Local Similarity 88.9%; Pred. No. 7.9;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 RPLPPLPXP 12
DB 2 RPLPPLPNP 10
RESULT 6
AAR93351
ID AAR93351 standard; peptide; 12 AA.
XX
AC AAR93351;
XX
XX 24-APR-1996 (first entry)
DT
XX FYN protein tyrosine kinase derived peptide #3.
DE
XX SH3 ligand; SH3 binding agent; biased phage library;
KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
KW p67; complex; chronic myelogenous leukaemia; cancer.
XX
OS Synthetic.
OS
XX WO9524419-A1.
PN
XX 14-SEP-1995.
PD

XX 13-MAR-1995; 95WO-US003208.
PF
XX 11-MAR-1994; 94US-00209835.
PR
XX 06-JAN-1995; 95US-00369832.
PR
XX (ARIA-) ARIAD PHARM INC.
PA
XX Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
PI
XX WPI; 1995-328231/42.
DR
XX Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.
PT
XX Disclosure; Fig 2; 74pp; English.
PS
XX The sequences given in AAR93343-68 represent peptides which are SH3
CC ligands/SH3 binding agents. They represent a biased phage library which
CC comprises six random amino acids flanking the hexapeptide RSLRPL- which
CC was identified as a recognition sequence for the src SH3 domain. These
CC sequences were identified using the method of the invention. The method
CC comprises contacting the SH3 domain with a mixture of peptides under
CC conditions permitting a ligand to bind to an SH3 domain to form a
CC complex. Any unbound peptides are removed and the complexed peptide
CC ligands are dissociated from the complexes. The selected peptides are
CC enriched by re-contacting them with the SH3 domain and then candidates
CC which bind to the SH3 domain are detected. The isolated SH3 binding
CC peptides may be used in the diagnosis, prevention and treatment of
CC conditions or diseases resulting from cellular processes mediated by an
CC SH3-based interaction. Such diseases include Paget's disease. Other
CC conditions treatable with these peptides include restenosis, rheumatoid
CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
CC p47 and p67 complex is implicated, etc
XX
SQ Sequence 12 AA;
Query Match 97.9%; Score 47; DB 2; Length 12;
Best Local Similarity 88.9%; Pred. No. 8.5;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 RPLPPLPXP 12
DB 4 RPLPPLPVP 12
RESULT 7
AAR93359
ID AAR93359 standard; peptide; 12 AA.
XX
AC AAR93359;
XX
XX 24-APR-1996 (first entry)
DT
XX LYN protein tyrosine kinase derived peptide #1.
DE
XX SH3 ligand; SH3 binding agent; biased phage library;
KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
KW p67; complex; chronic myelogenous leukaemia; cancer.
XX
OS Synthetic.
OS
XX WO9524419-A1.
PN
XX 14-SEP-1995.
PD
XX 13-MAR-1995; 95WO-US003208.
PF
XX 11-MAR-1994; 94US-00209835.
PR
XX 06-JAN-1995; 95US-00369832.
PR
XX

PS Claim 39; Page 304; 608pp; English.

XX The present invention describes composition of matter (I) comprising an

CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:

CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each

CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-

CC (L2)d-P2-(L3)e-P³, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,

CC (L2)d-P2-(L3)e-P³, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,

CC P3, and P4 = are each independently sequences of pharmacologically active

CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,

CC c, d, e, and f = are each independently 0 or 1, provided that at least 1

CC of a and b is 1. The composition can have cytostatic, antiasthmatic,

CC thrombolytic and immunosuppressive activities. DNAs, vectors and host

CC cells from the present invention can be used for producing pharmaceutical

CC compositions. The compositions are useful for treating cancer, asthma,

CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than

CC a Fab domain) can provide a longer half-life or incorporate functions

CC such as Fc receptor binding, protein A binding, complement fixation, and

CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to

CC AAB18003 represent nucleotide and amino acid sequences used in the

CC exemplification of the present invention

XX

SQ Sequence 12 AA;

Query Match 97.9%; Score 47; DB 3; Length 12;

Best Local Similarity 100.0%; Pred. No. 8.5;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RPLPPLPXP 12

Db |||||

4 RPLPPLPXP 12

RESULT 10

AAB17251

ID AAB17251 standard; peptide; 12 AA.

XX

AC AAB17251;

XX

DT 31-OCT-2000 (first entry)

XX

DE SH3 antagonist peptide sequence SEQ ID NO:307.

XX

KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;

KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;

KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;

KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;

KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;

KW vascular endothelial growth factor; matrix metalloproteinase; asthma;

XX thrombosis; pharmaceutical.

OS Synthetic.

XX

PN WO200024782-A2.

XX

PD 04-MAY-2000.

XX

PF 25-OCT-1999; 99WO-US025044.

XX

PR 23-OCT-1998; 98US-0105371P.

PR 22-OCT-1999; 99US-00428082.

XX

PA (AMGE-) AMGEN INC.

XX

PI Feige U, Liu C, Cheetham J, Boone TC;

XX

DR WPI; 2000-350702/30.

XX

PT Novel composition of matter comprising an Fc domain and pharmacologically

PT active peptides, useful for treating cancer and autoimmune diseases.

XX

PS Claim 39; Page 303; 608pp; English.

XX

CC The present invention describes composition of matter (I) comprising an

CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:

CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each

CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-

CC (L2)d-P2-(L3)e-P³, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,

CC P3, and P4 = are each independently sequences of pharmacologically active

CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,

CC c, d, e, and f = are each independently 0 or 1, provided that at least 1

CC of a and b is 1. The composition can have cytostatic, antiasthmatic,

CC thrombolytic and immunosuppressive activities. DNAs, vectors and host

CC cells from the present invention can be used for producing pharmaceutical

CC compositions. The compositions are useful for treating cancer, asthma,

CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than

CC a Fab domain) can provide a longer half-life or incorporate functions

CC such as Fc receptor binding, protein A binding, complement fixation, and

CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to

CC AAB18003 represent nucleotide and amino acid sequences used in the

CC exemplification of the present invention

XX

SQ Sequence 12 AA;

Query Match 97.9%; Score 47; DB 3; Length 12;

Best Local Similarity 100.0%; Pred. No. 8.5;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RPLPPLPXP 12

Db |||||

4 RPLPPLPXP 12

RESULT 11

ABB73244

ID ABB73244 standard; peptide; 12 AA.

XX

AC ABB73244;

XX

DT 05-APR-2002 (first entry)

XX

DE Src homology3 (SH3) antagonist peptide SEQ ID NO:307.

XX

KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;

KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;

KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;

KW TPO mimetic peptide; EPO mimetic peptide; BMP; VEGF antagonist;

KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;

KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;

KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;

KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;

KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;

KW sleep disorder; neurological degenerative disease; anaemia;

KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;

KW Fanconi's syndrome.

XX

OS Homo sapiens.

OS Synthetic.

XX

PN WO200183525-A2.

XX

PD 08-NOV-2001.

XX

PF 02-MAY-2001; 2001WO-US014310.

XX

PR 03-MAY-2000; 2000US-00563286.

XX

PA (AMGE-) AMGEN INC.

XX

PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;

XX

DR WPI; 2002-130313/17.

XX

PT Novel vehicle-peptide molecule or its multimers useful for treating

PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,

PT diabetic retinopathy, obesity, sleep disorders and infertility.

XX

"

PS	Claim 39; Page 55; 176pp; English.	XX	WPI; 2002-130313/17.
XX	The present invention describes a vehicle-peptide molecule (I) or its	XX	Novel vehicle-peptide molecule or its multimers useful for treating
CC	multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,	PT	inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
CC	cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,	PT	diabetic retinopathy, obesity, sleep disorders and infertility.
CC	antianaemic, anorectic, antiinfertility, haemostatic, dermatological and	XX	
CC	neuroprotective activities. (I) can be used as a therapeutic or	PS	Claim 39; Page 55; 176pp; English.
CC	prophylactic agent as well as for screening purposes. (I) is useful for	XX	
CC	diagnosing diseases characterised by dysfunction of their associated	CC	The present invention describes a vehicle-peptide molecule (I) or its
CC	protein of interest, for identifying normal or abnormal proteins of	CC	multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC	interest, as a part of diagnostic kit to detect the presence of their	CC	cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC	proteins of interest in a biological sample. Additionally, (I) is useful	CC	antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC	for treating inflammatory and autoimmune diseases, tumour growth, cancer,	CC	neuroprotective activities. (I) can be used as a therapeutic or
CC	rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,	CC	prophylactic agent as well as for screening purposes. (I) is useful for
CC	infertility, and neurological degenerative diseases. (I), comprising EPO-	CC	diagnosing diseases characterised by dysfunction of their associated
CC	mimetic compounds are useful for treating disorders characterised by low	CC	protein of interest, for identifying normal or abnormal proteins of
CC	red blood cell levels such as anaemia. The TPO-mimetic comprising	CC	interest, as a part of diagnostic kit to detect the presence of their
CC	compounds are useful for treating conditions that involve an existing	CC	proteins of interest in a biological sample. Additionally, (I) is useful
CC	megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet	CC	for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC	deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic	CC	rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC	tumour which result in thrombocytopaenia, systemic lupus erythematosus,	CC	infertility, and neurological degenerative diseases. (I), comprising EPO-
CC	and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777	CC	mimetic compounds are useful for treating disorders characterised by low
CC	represent amino acid and nucleic acid sequences used in the	CC	red blood cell levels such as anaemia. The TPO-mimetic comprising
CC	exemplification of the present invention	CC	compounds are useful for treating conditions that involve an existing
XX		CC	megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
SQ	Sequence 12 AA;	CC	deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
		CC	tumour which result in thrombocytopaenia, systemic lupus erythematosus,
		CC	and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
		CC	represent amino acid and nucleic acid sequences used in the
		CC	exemplification of the present invention
		XX	
QY	4 RPLPPLXP 12	SQ	Sequence 12 AA;
Db			
	4 RPLPPLXP 12		
RESULT 12			
ABB73247			
ID	ABB73247 standard; peptide; 12 AA.		
XX			
AC	ABB73247;		
XX			
DT	05-APR-2002 (first entry)		
XX			
DE	Src homology3 (SH3) antagonist peptide SEQ ID NO:310.		
XX			
KW	Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;		
KW	erythropoietin; TPO; tumour necrosis factor alpha inhibitor;		
KW	TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;		
KW	TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;		
KW	MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;		
KW	cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;		
KW	antianaemic; anorectic; antiinfertility; haemostatic; dermatological;		
KW	neuroprotective; inflammatory disease; autoimmune disease; tumour growth;		
KW	cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;		
KW	sleep disorder; neurological degenerative disease; anaemia;		
KW	thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;		
KW	Fanconi's syndrome.		
XX			
OS	Homo sapiens.		
OS	Synthetic.		
XX			
PN	WO200183525-A2.		
XX			
PD	08-NOV-2001.		
XX			
PF	02-MAY-2001; 2001WO-US014310.		
XX			
PR	03-MAY-2000; 2000US-00563286.		
XX			
PA	(AMGE-) AMGEN INC.		
XX			
PI	Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;		

XX	Query Match	97.9%;	Score 47;	DB 5;	Length 12;
DR	Best Local Similarity	100.0%;	Pred. No. 8.5;		
XX	Matches	9;	Conservative	0;	Mismatches
				0;	Indels
					0;
					Gaps
					0;
QY	4 RPLPPLXP 12				
Db					
	4 RPLPPLXP 12				
RESULT 13					
ADJ73401					
ID	ADJ73401 standard; peptide; 12 AA.				
XX					
AC	ADJ73401;				
XX					
DT	06-MAY-2004 (first entry)				
XX					
DE	SH3 antagonist peptide sequence SeqID 856.				
XX					
KW	mimetic; CDR mimetibody; gene therapy; transgenic; immune;				
KW	cardiovascular; infectious; malignant; neurologic disease; anaemia;				
KW	immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;				
KW	SH3.				
XX					
OS	Synthetic.				
XX					
PN	WO2003084477-A2.				
XX					
PD	16-OCT-2003.				
XX					
PF	24-MAR-2003; 2003WO-US009139.				
XX					
PR	29-MAR-2002; 2002US-0368791P.				
XX					
PA	(CENZ) CENTOCOR INC.				
XX					
PI	Heavner GA, Knight DM, Scallon BJ, Ghayeb J;				
XX					
DR	WPI; 2003-804237/75.				

XX New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.
XX
PS Disclosure; SEQ ID NO 856; 97pp; English.
XX
CC This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/ or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is an SH3 antagonist peptide sequence used to make a
CC mimetibody of the invention.
XX
SQ Sequence 12 AA;
Query Match 97.9%; Score 47; DB 7; Length 12;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 RPLPPLXP 12
Db |||||
4 RPLPPLXP 12
RESULT 14
ADJ73398
ID ADJ73398 standard; peptide; 12 AA.
XX
AC ADJ73398;
XX
DT 06-MAY-2004 (first entry)
XX
DE SH3 antagonist peptide sequence SeqID 853.
XX
KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW SH3.
XX
OS Synthetic.
XX
PN WO2003084477-A2.
XX
PD 16-OCT-2003.
XX
PF 24-MAR-2003; 2003WO-US009139.
XX
PR 29-MAR-2002; 2002US-0368791P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Scallion BJ, Ghayeb J;
XX
DR WPI; 2003-804237/75.
XX
PT New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.
XX
PS Disclosure; SEQ ID NO 853; 97pp; English.

XX This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/ or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is an SH3 antagonist peptide sequence used to make a
CC mimetibody of the invention.
XX
SQ Sequence 12 AA;
Query Match 97.9%; Score 47; DB 7; Length 12;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 RPLPPLXP 12
Db |||||
4 RPLPPLXP 12
RESULT 15
ADJ53032
ID ADJ53032 standard; peptide; 12 AA.
XX
AC ADJ53032;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID853.
XX
KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX
OS Unidentified.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1..12
FT /note= "All Xaa's in this sequence are unidentified amino
FT acids"
XX
PN WO2004002417-A2.
XX
PD 08-JAN-2004.
XX
PF 27-JUN-2003; 2003WO-US020347.
XX
PR 28-JUN-2002; 2002US-0392431P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Neaspor TC;
XX
PI Kutoloski KA;
XX
DR WPI; 2004-082870/08.
XX
PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious

PT diseases.
XX
PS Claim 3; SEQ ID NO 853; 129pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
XX used during the creation of a mimetibody of the invention.

SQ Sequence 12 AA;

Query Match 97.9%; Score 47; DB 8; Length 12;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
|||
Db 4 RPLPPLXP 12

RESULT 16
ADJ53035
ID ADJ53035 standard; peptide; 12 AA.

XX
AC ADJ53035;

XX
DT 06-MAY-2004 (first entry)

XX
DE CH1 deleted mimetibody-related peptide SeqID856.

XX
KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.

XX
OS Unidentified.
OS Synthetic.

XX
FH Key Location/Qualifiers

FT Misc-difference 1. .12
FT /note= "All Xaa's in this sequence are unidentified amino
FT acids"

XX
PN WO2004002417-A2.

XX
PD 08-JAN-2004.

XX
PF 27-JUN-2003; 2003WO-US020347.

XX
PR 28-JUN-2002; 2002US-0392431P.

XX
PA (CENZ) CENTOCOR INC.

XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallan BJ, Nesspor TC;
PI Kutoloski KA;

XX
DR WPI; 2004-082870/08.

XX
PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for

PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.

XX
PS Claim 3; SEQ ID NO 856; 129pp; English.

XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
XX used during the creation of a mimetibody of the invention.

SQ Sequence 12 AA;

Query Match 97.9%; Score 47; DB 8; Length 12;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
|||
Db 4 RPLPPLXP 12

RESULT 17

ADJ51993

ID ADJ51993 standard; peptide; 12 AA.

XX
AC ADJ51993;

XX
DT 06-MAY-2004 (first entry)

XX
DE CH1 deleted mimetibody-related peptide SeqID853.

XX
KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.

XX
OS Unidentified.
OS Synthetic.

XX
FH Key Location/Qualifiers

FT Misc-difference 1. .12
FT /note= "All Xaa's in this sequence are unidentified amino
FT acids"

XX
PN WO2004002424-A2.

XX
PD 08-JAN-2004.

XX
PF 30-JUN-2003; 2003WO-US020495.

XX
PR 28-JUN-2002; 2002US-0392431P.

PR 19-SEP-2002; 2002US-0412144P.
XX (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;
XX
XX WPI; 2004-082872/08.
XX
XX New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.
XX
XX Claim 15; SEQ ID NO 853; 123pp; English.
PS
XX This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CH1 deleted mimetibody, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetibody of the invention.
XX
SQ Sequence 12 AA;
Query Match 97.9%; Score 47; DB 8; Length 12;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 4 RPLPPLPXP 12
Db |||||||
4 RPLPPLPXP 12
RESULT 18
ADJ51996
ID ADJ51996 standard; peptide; 12 AA.
XX
AC ADJ51996;
XX
XX 06-MAY-2004 (first entry)
XX
XX CH1 deleted mimetibody-related peptide SeqID856.
XX
XX CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.
XX
OS Unidentified.

OS Synthetic.
XX
XX Key Location/Qualifiers
FH Misc-difference 1. .12
FT /note= "All Xaa's in this sequence are unidentified amino
FT acids"
XX
XX WO2004002424-A2.
XX
XX 08-JAN-2004.
XX
XX 30-JUN-2003; 2003WO-US020495.
XX
XX 28-JUN-2002; 2002US-0392431P.
PR 19-SEP-2002; 2002US-0412144P.
XX
XX (CENZ) CENTOCOR INC.
XX
XX Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;
XX
XX WPI; 2004-082872/08.
XX
XX New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.
XX
XX Claim 15; SEQ ID NO 856; 123pp; English.
PS
XX This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CH1 deleted mimetibody, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetibody of the invention.
XX
SQ Sequence 12 AA;
Query Match 97.9%; Score 47; DB 8; Length 12;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 4 RPLPPLPXP 12
Db |||||||
4 RPLPPLPXP 12
RESULT 19
AAR93461
ID AAR93461 standard; peptide; 15 AA.
XX
AC AAR93461;
XX
XX 09-MAY-1996 (first entry)
DT
XX
XX GST-PI3K protein tyrosine kinase derived peptide #1.
DE
XX SH3 ligand; SH3 binding agent; biased phage library;
KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
KW

KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
XX p67; complex; chronic myelogenous leukaemia; cancer.
OS Synthetic.
XX WO9524419-A1.
PN
XX
XX
PD 14-SEP-1995.
XX
XX 13-MAR-1995; 95WO-US003208.
PF
XX 11-MAR-1994; 94US-00209835.
PR 06-JAN-1995; 95US-00369832.
XX
XX (ARIA-) ARIAD PHARM INC.
PA
XX Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
PI WPI; 1995-328231/42.
XX
XX Identification of peptide(s) binding specifically to SH3 domains - for use in inhibiting interactions mediated by SH3 domains in treatment of e.g. osteoporosis and cancer.
PT
XX Disclosure; Fig 5; 74pp; English.
XX The sequences given in AAR93457-71 represent peptides which are SH3 ligands/SH3 binding agents. They represent a biased phage library which comprises five random amino acids flanking the decapeptide -RSLRPLPLP or derivatives of this, which was identified as a recognition sequence for the src SH3 domain. These sequences were identified using the method of the invention. The method comprises contacting the SH3 domain with a mixture of peptides under conditions permitting a ligand to bind to an SH3 domain to form a complex. Any unbound peptides are removed and the complexed peptide ligands are dissociated from the complexes. The selected peptides are enriched by re-contacting them with the SH3 domain and then candidates which bind to the SH3 domain are detected. The isolated SH3 binding peptides may be used in the diagnosis, prevention and treatment of conditions or diseases resulting from cellular processes mediated by an SH3-based interaction. Such diseases include Paget's disease. Other conditons treatable with these peptides include restenosis, rheumatoid arthritis, gout and other problems in which an SH3 of neutrophil oxidase p47 and p67 complex is implicated, etc
XX
SQ Sequence 15 AA;
Query Match 97.9%; Score 47; DB 2; Length 15;
Best Local Similarity 88.9%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 RPLPPLPXP 12
Db |||||
4 RPLPPLPFP 12
RESULT 20
AAR93457
ID AAR93457 standard; peptide; 15 AA.
XX
AC AAR93457;
XX
DT 09-MAY-1996 (first entry)
XX
DE GST-LYN protein tyrosine kinase derived peptide #1.
XX
KW SH3 ligand; SH3 binding agent; biased phage library;
KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
KW p67; complex; chronic myelogenous leukaemia; cancer.
XX
OS Synthetic.
XX
PN WO9524419-A1.

XX 14-SEP-1995.
XX
XX 13-MAR-1995; 95WO-US003208.
PF
XX 11-MAR-1994; 94US-00209835.
PR 06-JAN-1995; 95US-00369832.
XX
XX (ARIA-) ARIAD PHARM INC.
PA
XX Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
PI WPI; 1995-328231/42.
XX
XX Identification of peptide(s) binding specifically to SH3 domains - for use in inhibiting interactions mediated by SH3 domains in treatment of e.g. osteoporosis and cancer.
PT
XX Disclosure; Fig 5; 74pp; English.
XX The sequences given in AAR93457-71 represent peptides which are SH3 ligands/SH3 binding agents. They represent a biased phage library which comprises five random amino acids flanking the decapeptide -RSLRPLPLP or derivatives of this, which was identified as a recognition sequence for the src SH3 domain. These sequences were identified using the method of the invention. The method comprises contacting the SH3 domain with a mixture of peptides under conditions permitting a ligand to bind to an SH3 domain to form a complex. Any unbound peptides are removed and the complexed peptide ligands are dissociated from the complexes. The selected peptides are enriched by re-contacting them with the SH3 domain and then candidates which bind to the SH3 domain are detected. The isolated SH3 binding peptides may be used in the diagnosis, prevention and treatment of conditions or diseases resulting from cellular processes mediated by an SH3-based interaction. Such diseases include Paget's disease. Other conditons treatable with these peptides include restenosis, rheumatoid arthritis, gout and other problems in which an SH3 of neutrophil oxidase p47 and p67 complex is implicated, etc
XX
SQ Sequence 15 AA;
Query Match 97.9%; Score 47; DB 2; Length 15;
Best Local Similarity 88.9%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 RPLPPLPXP 12
Db |||||
4 RPLPPLPLP 12
RESULT 21
AAR93460
ID AAR93460 standard; peptide; 15 AA.
XX
AC AAR93460;
XX
DT 09-MAY-1996 (first entry)
XX
DE GST-LYN protein tyrosine kinase derived peptide #4.
XX
KW SH3 ligand; SH3 binding agent; biased phage library;
KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
KW p67; complex; chronic myelogenous leukaemia; cancer.
XX
OS Synthetic.
XX
PN WO9524419-A1.
XX
PD 14-SEP-1995.
XX
XX 13-MAR-1995; 95WO-US003208.
XX
PR 11-MAR-1994; 94US-00209835.

PR 06-JAN-1995; 95US-00369832.
XX
XX (ARIA-) ARIAD PHARM INC.
XX
PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
XX
DR WPI; 1995-328231/42.
XX
PT Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.
XX
PS Disclosure; Fig 5; 74pp; English.
XX
CC The sequences given in AAR93457-71 represent peptides which are SH3
CC ligands/SH3 binding agents. They represent a biased phage library which
CC comprises five random amino acids flanking the decapeptide -RSLRPLPLP or
CC derivatives of this, which was identified as a recognition sequence for
CC the src SH3 domain. These sequences were identified using the method of
CC the invention. The method comprises contacting the SH3 domain with a
CC mixture of peptides under conditions permitting a ligand to bind to an
CC SH3 domain to form a complex. Any unbound peptides are removed and the
CC complexed peptide ligands are dissociated from the complexes. The
CC selected peptide ligands are enriched by re-contacting them with the SH3 domain
CC and then candidates which bind to the SH3 domain are detected. The
CC isolated SH3 binding peptides may be used in the diagnosis, prevention
CC and treatment of conditions or diseases resulting from cellular processes
CC mediated by an SH3-based interaction. Such diseases include Paget's
CC disease. Other conditions treatable with these peptides include
CC restenosis, rheumatoid arthritis, gout and other problems in which an SH3
CC of neutrophil oxidase p47 and p67 complex is implicated, etc
XX
SQ Sequence 15 AA;

Query Match 97.9%; Score 47; DB 2; Length 15;
Best Local Similarity 88.9%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLPXP 12
Db |||||||
4 RPLPPLPLP 12

RESULT 22
AAR93458
ID AAR93458 standard; peptide; 15 AA.
XX
AC AAR93458;
XX
DT 09-MAY-1996 (first entry)
XX
DE GST-LYN protein tyrosine kinase derived peptide #2.
XX
KW SH3 ligand; SH3 binding agent; biased phage library;
KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
KW p67; complex; chronic myelogenous leukaemia; cancer.
XX
OS Synthetic.
XX
PN WO9524419-A1.
XX
PD 14-SEP-1995.
XX
PF 13-MAR-1995; 95WO-US003208.
XX
PR 11-MAR-1994; 94US-00209835.
PR 06-JAN-1995; 95US-00369832.
XX
PA (ARIA-) ARIAD PHARM INC.
XX
PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
XX

DR WPI; 1995-328231/42.
XX
PT Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.
XX
PS Disclosure; Fig 5; 74pp; English.
XX
CC The sequences given in AAR93457-71 represent peptides which are SH3
CC ligands/SH3 binding agents. They represent a biased phage library which
CC comprises five random amino acids flanking the decapeptide -RSLRPLPLP or
CC derivatives of this, which was identified as a recognition sequence for
CC the src SH3 domain. These sequences were identified using the method of
CC the invention. The method comprises contacting the SH3 domain with a
CC mixture of peptides under conditions permitting a ligand to bind to an
CC SH3 domain to form a complex. Any unbound peptides are removed and the
CC complexed peptide ligands are dissociated from the complexes. The
CC selected peptide ligands are enriched by re-contacting them with the SH3 domain
CC and then candidates which bind to the SH3 domain are detected. The
CC isolated SH3 binding peptides may be used in the diagnosis, prevention
CC and treatment of conditions or diseases resulting from cellular processes
CC mediated by an SH3-based interaction. Such diseases include Paget's
CC disease. Other conditions treatable with these peptides include
CC restenosis, rheumatoid arthritis, gout and other problems in which an SH3
CC of neutrophil oxidase p47 and p67 complex is implicated, etc
XX
SQ Sequence 15 AA;

Query Match 97.9%; Score 47; DB 2; Length 15;
Best Local Similarity 88.9%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLPXP 12
Db |||||||
4 RPLPPLPLP 12

RESULT 23
AAR93459
ID AAR93459 standard; peptide; 15 AA.
XX
AC AAR93459;
XX
DT 09-MAY-1996 (first entry)
XX
DE GST-LYN protein tyrosine kinase derived peptide #3.
XX
KW SH3 ligand; SH3 binding agent; biased phage library;
KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
KW p67; complex; chronic myelogenous leukaemia; cancer.
XX
OS Synthetic.
XX
PN WO9524419-A1.
XX
PD 14-SEP-1995.
XX
PF 13-MAR-1995; 95WO-US003208.
XX
PR 11-MAR-1994; 94US-00209835.
PR 06-JAN-1995; 95US-00369832.
XX
PA (ARIA-) ARIAD PHARM INC.
XX
PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
XX
DR WPI; 1995-328231/42.
XX
PT Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.
XX

PS Disclosure; Fig 5; 74pp; English.

XX The sequences given in AAR93457-71 represent peptides which are SH3

CC ligands/SH3 binding agents. They represent a biased phage library which

CC comprises five random amino acids flanking the decapeptide -RSLRPLPLP or

CC derivatives of this, which was identified as a recognition sequence for

CC the src SH3 domain. These sequences were identified using the method of

CC the invention. The method comprises contacting the SH3 domain with a

CC mixture of peptides under conditions permitting a ligand to bind to an

CC SH3 domain to form a complex. Any unbound peptides are removed and the

CC complexed peptide ligands are dissociated from the complexes. The

CC selected peptides are enriched by re-contacting them with the SH3 domain

CC and then candidates which bind to the SH3 domain are detected. The

CC isolated SH3 binding peptides may be used in the diagnosis, prevention

CC and treatment of conditions or diseases resulting from cellular processes

CC mediated by an SH3-based interaction. Such diseases include Paget's

CC disease. Other conditions treatable with these peptides include

CC restenosis, rheumatoid arthritis, gout and other problems in which an SH3

CC of neutrophil oxidase p47 and p67 complex is implicated, etc

XX Sequence 15 AA;

SQ Query Match 97.9%; Score 47; DB 2; Length 15;

Best Local Similarity 88.9%; Pred. No. 10;

Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12

|||||||

Db 4 RPLPPLPLP 12

RESULT 24

AAW05451

ID AAW05451 standard; peptide; 15 AA.

XX AAW05451;

DT 24-FEB-1998 (first entry)

XX SH3-binding peptide bSH3002.

DE Src-homology region 3 domain; human; mouse; SH3 domain; cell growth;

XX cellular signalling element; cellular structural element; malignancy;

KW protein identification; functional domain; protein screening;

KW cellular signal transduction process; binding peptide.

XX Synthetic.

OS WO9631625-A1.

PN 10-OCT-1996.

XX 04-APR-1996; 96WO-US004454.

XX 07-APR-1995; 95US-00417872.

PR 03-APR-1996; 96US-00630915.

XX (CYTO-) CYTOGEN CORP.

PA (UYNC-) UNIV NORTH CAROLINA.

XX Sparks AB, Hoffman N, Kay BK, Fowlkes DM, McConnell SJ;

XX WPI; 1996-465045/46.

XX Identifying polypeptide(s) having specific functional domain (esp. SH3

PT domain) - comprises detecting selective binding to recognition unit,

PT regardless of sequence homology.

XX Example; Fig 12A; 174pp; English.

XX AAW05445-W05492 represent Src-homology region 3 (SH2) domain binding

CC peptides. These sequences were used as parts of multivalent recognition

CC unit complexes used in the method of the invention. The method of the

CC invention is for identifying polypeptides containing functional domains

CC of interest (especially SH3 domains). It comprises contacting a

CC multivalent recognition unit (RU) complex with a number of peptides and

CC identifying polypeptides having a selective binding affinity for the RU

CC complex. The method is based on functional similarities and does not rely

CC on sequence similarities. Prior methods only gave limited success for

CC identifying proteins containing an SH3 domain due to the minimal sequence

CC homology among known SH3 proteins. Multivalent RU complexes are

CC particularly suited to screening for polypeptides containing functional

CC domains that are similar to, but not identical in sequence to, the

CC original target functional domain. The new method enables proteins having

CC a common function to be identified. Identification of novel SH3 proteins

CC will be useful for a better understanding of cell growth, malignancy,

CC signal transduction processes, etc. New candidate drugs can be

CC identified, and their specificities (e.g. pharmacological activities) can

CC be assessed using the method of the invention

XX Sequence 15 AA;

SQ Query Match 97.9%; Score 47; DB 2; Length 15;

Best Local Similarity 88.9%; Pred. No. 10;

Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12

|||||||

Db 5 RPLPPLPDP 13

RESULT 25

AAW37659

ID AAW37659 standard; peptide; 15 AA.

XX AAW37659;

DT 23-APR-1998 (first entry)

XX PPPPY motif containing peptide bSH3002 used to bind WW domains.

DE Peptide recognition unit; WW domain; cell signalling; growth regulation;

KW cytoskeleton organisation; targeted drug screening; modulator;

KW WW domain interaction; YAP protein; dystrophin.

XX Synthetic.

OS WO9737223-A1.

PN 09-OCT-1997.

XX 03-APR-1997; 97WO-US005547.

XX 03-APR-1996; 96US-00630916.

XX (CYTO-) CYTOGEN CORP.

PA (UYNC-) UNIV NORTH CAROLINA.

XX Pirozzi G, Kay BK, Fowlkes DM;

XX WPI; 1997-503234/46.

XX Identifying cell signalling and growth regulatory polypeptides by

PT reaction with multivalent recognition complex - polypeptides are useful

PT in targeted drug selection.

XX Example 6.3; Fig 7; 220pp; English.

XX Peptides AAW37653-77 contain ppppy-like motifs. The pppp motif is found

CC in the proline rich regions of WBP-1 and WBP-2 proteins. Peptides

CC containing this residue have been shown to bind the YAP WW domain, but

CC not the WW domain from dystrophin or to a panel of SH3 domains. Peptides

CC AAW37653-77 were biotinylated and complexed with alkaline streptavidin,

CC and used in a cross affinity mapping experiment. They were tested for

CC their ability to bind to the 12 individual novel WW domains of WWP1

CC (AAW36794), WWP2 (AAW36795), WWP3 (AAW37696) and WWP4 (AAW36797), which

CC were expressed as glutathione-S-transferase expression proteins. The
CC present peptide, derived from a potassium channel, does not bind to WW
CC domains of the novel proteins. The WW domain is a small functional
CC domain. Its name is derived from the observation that two tryptophan
CC residues, one in the amino terminal portion of the WW domain and one in
CC the carboxyl terminal portion, are conserved. Most proteins containing WW
CC domains have a function involving cell signalling and growth regulation
CC or the organisation of the cytoskeleton. Polypeptides containing a WW
CC domain are identified by treating a multivalent recognition unit complex
CC that has selective binding affinity for a WW domain, with many
CC polypeptides and identifying those with selective affinity for the
CC complex. Proteins containing WW domains are used for targeted drug
CC screening, i.e. to identify potential modulators of specific WW domain
CC interactions
XX
SQ Sequence 15 AA;

Query Match 97.9%; Score 47; DB 2; Length 15;
Best Local Similarity 88.9%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLPXP 12
Db 5 RPLPPLPDP 13
| | | | | | | |
RESULT 26
AAW38924
ID AAW38924 standard; peptide; 15 AA.
XX
AC AAW38924;
XX
DT 27-MAR-1998 (first entry)
XX
DE Peptide resembling an SH3 domain binding peptide SEQ ID NO:321.
XX
KW Cortactin; SH3 domain; binding peptide; Src homology region 3;
KW tyrosine kinase; immune response; lymphokine; interleukin 1; Nck; Abl;
KW PLCgamma; p53bp2; Crk; Yes; Grb2.
XX
OS Synthetic.
XX
PN WO9730074-A1.
XX
PD 21-AUG-1997.
XX
PF 14-FEB-1997; 97WO-US002298.
XX
PR 16-FEB-1996; 96US-00602999.
XX
PA (CYTO-) CYTOGEN CORP.
PA (UYNC-) UNIV NORTH CAROLINA.
XX
PI Sparks AB, Kay BK, Thorn JM, Quilliam LA, Der CJ, Fowlkes DM;
PI Rider JE;
XX
DR WPI; 1997-424972/39.
XX
PT Src homology region 3 binding peptide - used to activate Src tyrosine
PT kinase(s) and to stimulate immune response by increasing production of
PT certain lymphokine(s), e.g. interleukin-1.
XX
PS Claim 22; Page 90; 131pp; English.
XX
CC The present sequence represents a peptide which resembles a Src homology
CC region 3 (SH3) binding peptide. SH3 binding peptides are selected from:
CC (a) peptides which bind the SH3 domain of Cortactin; (b) peptides which
CC bind the middle SH3 domain of Nck; (c) peptides which bind the SH3 domain
CC of Abl; (d) peptides which bind the SH3 domain of Src; (e) peptides which
CC bind the SH3 domain of PLC gamma; (f) peptides which bind the SH3 domain
CC of p53bp2; (g) peptides which bind the amino-terminal SH3 domain of Crk;
CC (h) peptides which bind the SH3 domain of Yes; and (i) peptides which
CC bind the amino-terminal SH3 domain of Grb2. The purified binding peptides

CC can be used in the method to identify inhibitors of their binding to
CC their respective SH3 domains, which could be used to modulate the
CC pharmacological activity of proteins or polypeptide containing the SH3
CC domain. The peptides can also be used to activate Src or Src-related
CC protein tyrosine kinases, to stimulate the immune response by increasing
CC the production of certain lymphokines, e.g. tumour necrosis factor-alpha
CC and interleukin-1, or to deliver a conjugated molecule to certain
CC cellular compartments containing Src or Src related proteins
XX
SQ Sequence 15 AA;

Query Match 97.9%; Score 47; DB 2; Length 15;
Best Local Similarity 88.9%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLPXP 12
Db 5 RPLPPLPDP 13
| | | | | | | |
RESULT 27
ADB49285
ID ADB49285 standard; peptide; 15 AA.
XX
AC ADB49285;
XX
DT 04-DEC-2003 (first entry)
XX
DE Novel WW domain binding peptide #7.
XX
KW WW domain; drug candidate screening; drug discovery; drug modification;
KW drug refinement; immunogen; WW binding protein; WW domain.
XX
OS Unidentified.
XX
PN US2003077577-A1.
XX
PD 24-APR-2003.
XX
PF 28-JUN-2002; 2002US-00185050.
PR 03-APR-1996; 96US-00630916.
PR 03-APR-1997; 97US-00826516.
XX
PA (PIRO/) PIROZZI G.
PA (KAYB/) KAY B K.
PA (FOWL/) FOWLKES D M.
XX
PI Pirozzi G, Kay BK, Fowlkes DM;
XX
DR WPI; 2003-635075/60.
XX
PT Novel purified polypeptide comprising WW domain, useful for drug
PT discovery, modification and refinement, for discovering polypeptides
PT involved in pharmacological activities, or as an immunogen to generate
PT antibodies.
XX
PS Example; Fig 7; 133pp; English.
XX
CC The invention describes a purified polypeptide (I) comprising a WW domain
CC which has a sequence (S1) selected from 11 sequences fully defined in the
CC specification, a sequence (S2) selected from 48 sequences fully defined
CC in the specification or a sequence (S3) comprising 683, 906, 224 or 725
CC amino acids fully defined in the specification. (I) is useful for
CC screening a potential drug candidate, by allowing (I) to come into
CC contact with at least one recognition unit having a selective affinity
CC for the WW domain in (I), in the presence of an amount of a potential
CC drug candidate, such that (I) and the recognition unit are capable of
CC interacting when brought into contact with one another in the absence of
CC the drug candidate, and determining the effect, if any, of the presence
CC of the amount of the drug candidate on the interaction of (I) with the
CC recognition unit. (I) is useful for drug discovery, modification and
CC refinement, for discovering polypeptides involved in pharmacological

CC activities, or as an immunogen to generate antibodies. This is the amino
CC acid sequence of a peptide that binds the novel WW domains of the
CC invention
XX
SQ Sequence 15 AA;
Query Match 97.9%; Score 47; DB 7; Length 15;
Best Local Similarity 88.9%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 4 RPLPPLXP 12
Db 5 RPLPPLDP 13
RESULT 28
AAB17217
ID AAB17217 standard; peptide; 17 AA.
XX
AC AAB17217;
XX
DT 31-OCT-2000 (first entry)
XX
DE Mast cell antagonist/Mast cell protease inhibitor peptide SEQ ID NO:273.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
XX
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 39; Page 292; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-P1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention

XX
SQ Sequence 17 AA;
Query Match 97.9%; Score 47; DB 3; Length 17;
Best Local Similarity 88.9%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 4 RPLPPLXP 12
Db 6 RPLPPLPLP 14
RESULT 29
ABB73208
ID ABB73208 standard; peptide; 17 AA.
XX
AC ABB73208;
XX
DT 05-APR-2002 (first entry)
XX
DE Mast cell antagonist/mast cell protease inhibitor peptide SEQ ID NO:273.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 39; Page 54; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders, EPO-
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising

CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention

XX SQ Sequence 17 AA;

Query Match 97.9%; Score 47; DB 5; Length 17;
Best Local Similarity 88.9%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
Db 6 RPLPPLPLP 14

RESULT 30
ADJ73362 standard; peptide; 17 AA.

XX ADJ73362;
AC ADJ73362;
DT 06-MAY-2004 (first entry)
DE Mast cell antagonist peptide sequence SeqID 816.
XX mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurogenic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW mast cell.
XX Synthetic.
OS WO2003084477-A2.
PN 16-OCT-2003.
XX 24-MAR-2003; 2003WO-US009139.
PF 29-MAR-2002; 2002US-0368791P.
XX (CENZ) CENTOCOR INC.
PI Heavner GA, Knight DM, Scallion BJ, Ghrayeb J;
XX WPI; 2003-804237/75.
DR New CDR mimetibody comprising a portion of a heavy or light chain
XX variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.
XX Disclosure; SEQ ID NO 816; 97pp; English.

PS This invention relates to novel mammalian CDR mimetibodies, specific
XX portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/ or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is a mast cell antagonist peptide sequence used to make
CC a mimetibody of the invention.

XX SQ Sequence 17 AA;

Query Match 97.9%; Score 47; DB 7; Length 17;
Best Local Similarity 88.9%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
Db 6 RPLPPLPLP 14

RESULT 31
ADJ52996 standard; peptide; 17 AA.

XX ADJ52996;
AC ADJ52996;
DT 06-MAY-2004 (first entry)
XX CH1 deleted mimetibody-related peptide SeqID816.
DE CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX Unidentified.
OS Synthetic.
XX WO2004002417-A2.
PN 08-JAN-2004.
XX 27-JUN-2003; 2003WO-US020347.
PF 28-JUN-2002; 2002US-0392431P.
XX (CENZ) CENTOCOR INC.
XX Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;
XX WPI; 2004-082870/08.
DR New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
XX modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX Claim 3; SEQ ID NO 816; 129pp; English.

PS This invention relates to CH1 deleted mimetibodies (and the DNA sequences
XX which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
XX used during the creation of a mimetibody of the invention.

SQ Sequence 17 AA;

Query Match 97.9%; Score 47; DB 8; Length 17;
Best Local Similarity 88.9%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

QY 4 RPLPPLPXP 12
Db 6 RPLPPLPLP 14

RESULT 32
ADJ51957
ID ADJ51957 standard; peptide; 17 AA.
XX
AC ADJ51957;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID816.
XX
KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.
XX
OS Unidentified.
OS Synthetic.
XX
PN WO2004002424-A2.
XX
PD 08-JAN-2004.
XX
PF 30-JUN-2003; 2003WO-US020495.
XX
PR 28-JUN-2002; 2002US-0392431P.
PR 19-SEP-2002; 2002US-0412144P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallon BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082872/08.
XX
PT New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.
XX
PS Claim 15; SEQ ID NO 816; 123pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulatory, antiallergic, muscular-Gen, cytostatic,
CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CH1 deleted mimetibody, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,

CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetibody of the invention.
XX
SQ Sequence 17 AA;

Query Match 97.9%; Score 47; DB 8; Length 17;
Best Local Similarity 88.9%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLPXP 12
Db 6 RPLPPLPLP 14

RESULT 33
AAY41645
ID AAY41645 standard; peptide; 22 AA.
XX
AC AAY41645;
XX 02-DEC-1999 (first entry)
XX Human peptide hKv1.5-62-83.
DE SH3 domain; binding motif; potassium channel; protein tyrosine kinase;
KW proline rich.
XX Homo sapiens.
OS
XX US5955259-A.
XX 21-SEP-1999.
PF 19-DEC-1996; 96US-00769745.
XX
PR 19-DEC-1996; 96US-00769745.
XX
PA (UYER-) UNIV BRANDEIS.
XX
PI Holmes TC, Levitan IB;
XX
DR WPI; 1999-560490/47.
XX
PT Identification of compounds that modulate potassium ion channel binding
PT with protein tyrosine kinase SH3 domains.
XX
PS Example 2; Col 14; 18pp; English.
XX
CC A method has been developed for determining if a compound modulates the
CC binding of a potassium ion channel to the SH3 domain of a protein
CC tyrosine kinase by contacting the channel with a polypeptide comprising
CC the SH3 domain and the compound to be assessed and measuring channel-SH3
CC binding. The method is useful for assessing the ability of a compound to
CC modulate the formation of channel-SH3 domain complexes to improve the
CC understanding of mechanisms of potassium channel blockage and asses the
CC ability of potential therapeutics to inhibit blockage. The present
CC sequence represents the Human peptide hKv1.5-62-83, which is used in an
CC example from the present invention
XX
SQ Sequence 22 AA;

Query Match 97.9%; Score 47; DB 2; Length 22;
Best Local Similarity 88.9%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLPXP 12
Db 4 RPLPPLPDP 12

RESULT 34
AAAY41612
ID AAY41612 standard; peptide; 25 AA.
XX
XX AC AAY41612;
AC
XX 02-DEC-1999 (first entry)
DT
XX Mammalian ion channel proline rich motif containing peptide #6.
DE
XX SH3 domain; binding motif; potassium channel; protein tyrosine kinase;
KW proline rich.
XX
XX Homo sapiens.
OS
XX US5955259-A.
PN
XX 21-SEP-1999.
PD
XX 19-DEC-1996; 96US-00769745.
PF
XX 19-DEC-1996; 96US-00769745.
PR
XX (UYBR-) UNIV BRANDEIS.
PA
XX Holmes TC, Levitan IB;
PI
XX WPI; 1999-560490/47.
DR
XX Identification of compounds that modulate potassium ion channel binding
PT with protein tyrosine kinase SH3 domains.
PT
XX Disclosure; Col 8; 18pp; English.
PS
XX A method has been developed for determining if a compound modulates the
CC binding of a potassium ion channel to the SH3 domain of a protein
CC tyrosine kinase by contacting the channel with a polypeptide comprising
CC the SH3 domain and the compound to be assessed and measuring channel-SH3
CC binding. The method is useful for assessing the ability of a compound to
CC modulate the formation of channel-SH3 domain complexes to improve the
CC understanding of mechanisms of potassium channel blockage and asses the
CC ability of potential therapeutics to inhibit blockage. AAY41607 to
CC AAY41644 represent mammalian ion channel peptides with proline-rich
CC motifs
XX
SQ Sequence 25 AA;

Query Match 97.9%; Score 47; DB 2; Length 25;
Best Local Similarity 88.9%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
Db 1 RPLPPLPDP 9

RESULT 35
ABR42419
ID ABR42419 standard; protein; 613 AA.
XX
XX ABR42419;
AC
XX 11-AUG-2003 (first entry)
DT
XX Human potassium channel Kv1.5.
DE
XX Human; potassium channel; Kv1.5; neuronal disease; multiple sclerosis;
KW neuroprotective.
XX
XX Homo sapiens.
OS
XX WO2003035690-A2.
PN

XX 01-MAY-2003.
PD
XX 25-OCT-2002; 2002WO-GB004821.
PF
XX 25-OCT-2001; 2001GB-00025636.
PR
XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
PA
XX Dolly JO, Akhtar S, Shamotienko O;
PI
XX WPI; 2003-421404/39.
DR
XX Preparing a K+ channel fusion protein comprising alpha subunits and beta
PT subunits, useful in the manufacture of a medicament for treating multiple
PT sclerosis.
PT
XX Disclosure; Fig 4; 93pp; English.
PS
XX The present sequence is the protein sequence of human potassium channel
CC Kv1.5. The invention provides oligomeric potassium channel fusion
CC proteins, preferably comprising Kv1.1 and/or Kv1.2 subunits, and methods
CC of producing them in mammalian cells. The fusion proteins can be used to
CC identify compounds that modulate potassium channel activity. Such
CC modulator compounds are useful for the treatment of multiple sclerosis
CC
XX Sequence 613 AA;
SQ

Query Match 97.9%; Score 47; DB 6; Length 613;
Best Local Similarity 88.9%; Pred. No. 2.9e+02;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
Db 65 RPLPPLPDP 73

RESULT 36
AAR93362
ID AAR93362 standard; peptide; 12 AA.
XX
AC AAR93362;
XX
DT 24-APR-1996 (first entry)
DT
XX LYN/PI3K protein tyrosine kinase derived peptide.
DE
XX SH3 ligand; SH3 binding agent; biased phage library;
KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
KW p67; complex; chronic myelogenous leukaemia; cancer.
XX
OS Synthetic.
OS
XX WO9524419-A1.
PN
XX 14-SEP-1995.
PD
XX 13-MAR-1995; 95WO-US003208.
PF
XX 11-MAR-1994; 94US-00209835.
PR
XX 06-JAN-1995; 95US-00369832.
PR
XX (ARIA-) ARIAD PHARM INC.
PA
XX Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
XX
DR WPI; 1995-328231/42.
DR
XX Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.
PT
XX

PS Disclosure; Fig 2; 74pp; English.

XX The sequences given in AAR93343-68 represent peptides which are SH3

CC ligands/SH3 binding agents. They represent a biased phage library which

CC comprises six random amino acids flanking the hexapeptide RSLRPL- which

CC was identified as a recognition sequence for the src SH3 domain. These

CC sequences were identified using the method of the invention. The method

CC comprises contacting the SH3 domain with a mixture of peptides under

CC conditions permitting a ligand to bind to an SH3 domain to form a

CC complex. Any unbound peptides are removed and the complexed peptide

CC ligands are dissociated from the complexes. The selected peptides are

CC enriched by re-contacting them with the SH3 domain and then candidates

CC which bind to the SH3 domain are detected. The isolated SH3 binding

CC peptides may be used in the diagnosis, prevention and treatment of

CC conditions or diseases resulting from cellular processes mediated by an

CC SH3-based interaction. Such diseases include Paget's disease. Other

CC conditions treatable with these peptides include restenosis, rheumatoid

CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase

CC p47 and p67 complex is implicated, etc

XX

SQ Sequence 12 AA;

Query Match 95.8%; Score 46; DB 2; Length 12;

Best Local Similarity 88.9%; Pred. No. 11;

Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12

Db 4 RPLPPLPPP 12

RESULT 37

AAR93377

ID AAR93377 standard; peptide; 12 AA.

XX

AC AAR93377;

XX

DT 24-APR-1996 (first entry)

XX

DE Grb-2 protein tyrosine kinase derived peptide #1.

XX

KW SH3 ligand; SH3 binding agent; biased phage library;

KW recognition sequence; src SH3 domain; Paget's disease; restenosis;

KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;

KW p67; complex; chronic myelogenous leukaemia; cancer.

XX

OS Synthetic.

XX

PN WO9524419-A1.

XX

PD 14-SEP-1995.

XX

PF 13-MAR-1995; 95WO-US003208.

XX

PR 11-MAR-1994; 94US-00209835.

PR 06-JAN-1995; 95US-00369832.

XX

PA (ARIA-) ARIAD PHARM INC.

XX

PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;

XX

DR WPI; 1995-328231/42.

XX

XX Identification of peptide(s) binding specifically to SH3 domains - for

PT use in inhibiting interactions mediated by SH3 domains in treatment of

PT e.g. osteoporosis and cancer.

XX

PS Disclosure; Fig 2; 74pp; English.

XX

CC The sequences given in AAR93377-80 represent peptides which are SH3

CC ligands/SH3 binding agents. They represent a biased phage library which

CC comprises six random amino acids flanking the hexapeptide RLYRPL- which

CC was identified as a recognition sequence for the src SH3 domain. These

CC sequences were identified using the method of the invention. The method

CC comprises contacting the SH3 domain with a mixture of peptides under

CC conditions permitting a ligand to bind to an SH3 domain to form a

CC complex. Any unbound peptides are removed and the complexed peptide

CC ligands are dissociated from the complexes. The selected peptides are

CC enriched by re-contacting them with the SH3 domain and then candidates

CC which bind to the SH3 domain are detected. The isolated SH3 binding

CC peptides may be used in the diagnosis, prevention and treatment of

CC conditions or diseases resulting from cellular processes mediated by an

CC SH3-based interaction. Such diseases include Paget's disease. Other

CC conditions treatable with these peptides include restenosis, rheumatoid

CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase

CC p47 and p67 complex is implicated, etc

XX

SQ Sequence 12 AA;

Query Match 95.8%; Score 46; DB 2; Length 12;

Best Local Similarity 88.9%; Pred. No. 11;

Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12

Db 4 RPLPPLPWP 12

RESULT 38

AAR93345

ID AAR93345 standard; peptide; 12 AA.

XX

AC AAR93345;

XX

DT 24-APR-1996 (first entry)

XX

DE SRC protein tyrosine kinase derived peptide #3.

DE

XX

KW SH3 ligand; SH3 binding agent; biased phage library;

KW recognition sequence; src SH3 domain; Paget's disease; restenosis;

KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;

KW p67; complex; chronic myelogenous leukaemia; cancer.

XX

OS Synthetic.

XX

PN WO9524419-A1.

XX

PD 14-SEP-1995.

XX

PF 13-MAR-1995; 95WO-US003208.

XX

PR 11-MAR-1994; 94US-00209835.

PR 06-JAN-1995; 95US-00369832.

XX

PA (ARIA-) ARIAD PHARM INC.

XX

PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;

XX

DR WPI; 1995-328231/42.

XX

XX Identification of peptide(s) binding specifically to SH3 domains - for

PT use in inhibiting interactions mediated by SH3 domains in treatment of

PT e.g. osteoporosis and cancer.

XX

PS Disclosure; Fig 2; 74pp; English.

XX

CC The sequences given in AAR93343-68 represent peptides which are SH3

CC ligands/SH3 binding agents. They represent a biased phage library which

CC comprises six random amino acids flanking the hexapeptide RSLRPL- which

CC was identified as a recognition sequence for the src SH3 domain. These

CC sequences were identified using the method of the invention. The method

CC comprises contacting the SH3 domain with a mixture of peptides under

CC conditions permitting a ligand to bind to an SH3 domain to form a

CC complex. Any unbound peptides are removed and the complexed peptide

CC ligands are dissociated from the complexes. The selected peptides are

CC enriched by re-contacting them with the SH3 domain and then candidates

CC which bind to the SH3 domain are detected. The isolated SH3 binding
CC peptides may be used in the diagnosis, prevention and treatment of
CC conditions or diseases resulting from cellular processes mediated by an
CC SH3-based interaction. Such diseases include Paget's disease. Other
CC conditions treatable with these peptides include restenosis, rheumatoid
CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
CC p47 and p67 complex is implicated, etc

XX Sequence 12 AA;

Query Match 95.8%; Score 46; DB 2; Length 12;
Best Local Similarity 88.9%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
| | | | | | |
Db 4 RPLPPLPPP 12

RESULT 39
AAB17255
ID AAB17255 standard; peptide; 12 AA.
XX
AC AAB17255;
XX
DT 31-OCT-2000 (first entry)
XX
DE SH3 antagonist peptide sequence SEQ ID NO:311.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 39; Page 304; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a, and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions

CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention

XX Sequence 12 AA;

Query Match 95.8%; Score 46; DB 3; Length 12;
Best Local Similarity 88.9%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
| | | | | | |
Db 4 RPLPPLPPP 12

RESULT 40
ABB73248
ID ABB73248 standard; peptide; 12 AA.
XX
AC ABB73248;
XX
DT 05-APR-2002 (first entry)
XX
DE Src homology3 (SH3) antagonist peptide SEQ ID NO:311.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IGG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 39; Page 55; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,

CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 12 AA;

Query Match 95.8%; Score 46; DB 5; Length 12;
Best Local Similarity 88.9%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLPPX 12
Db ||||| |
4 RPLPPLPPP 12

RESULT 41
ADJ73402
ID ADJ73402 standard; peptide; 12 AA.
XX
AC ADJ73402;
XX
DT 06-MAY-2004 (first entry)
XX
DE SH3 antagonist peptide sequence SeqID 857.
XX
KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW SH3.
XX
OS Synthetic.
XX
XX WO2003084477-A2.
PN
XX
PD 16-OCT-2003.
XX
PF 24-MAR-2003; 2003WO-US009139.
XX
PR 29-MAR-2002; 2002US-0368791P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Scallion BJ, Ghrayeb J;
XX
DR WPI; 2003-804237/75.
XX
PT New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.
XX
PS Disclosure; SEQ ID NO 857; 97pp; English.
XX
CC This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/ or neurologic

CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is an SH3 antagonist peptide sequence used to make a
CC mimetibody of the invention.
XX
SQ Sequence 12 AA;

Query Match 95.8%; Score 46; DB 7; Length 12;
Best Local Similarity 88.9%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLPPX 12
Db ||||| |
4 RPLPPLPPP 12

RESULT 42
ADJ53036
ID ADJ53036 standard; peptide; 12 AA.
XX
AC ADJ53036;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID857.
XX
KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX
OS Unidentified.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1..12
FT /note= "All Xaa's in this sequence are unidentified amino
FT acids"
XX
PN WO2004002417-A2.
XX
PD 08-JAN-2004.
XX
PF 27-JUN-2003; 2003WO-US020347.
XX
PR 28-JUN-2002; 2002US-0392431P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082870/08.
XX
PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX
PS Claim 3; SEQ ID NO 857; 129pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an

CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.

XX
SQ Sequence 12 AA;

Query Match 95.8%; Score 46; DB 8; Length 12;
Best Local Similarity 88.9%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
| | | | | | |
Db 4 RPLPPLPPP 12

RESULT 43
ADJ51997
ID ADJ51997 standard; peptide; 12 AA.
XX
AC ADJ51997;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID857.
XX
KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.

XX
OS Unidentified.
OS Synthetic.

XX
FH Key Location/Qualifiers
FT Misc-difference 1. .12
FT /note= "All Xaa's in this sequence are unidentified amino
FT acids"
XX
PN WO2004002424-A2.
XX
PD 08-JAN-2004.
XX
PF 30-JUN-2003; 2003WO-US020495.
XX
PR 28-JUN-2002; 2002US-0392431P.
PR 19-SEP-2002; 2002US-0412144P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082872/08.
XX
PT New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.

XX
PS Claim 15; SEQ ID NO 857; 123pp; English.

XX This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF) -
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CH1 deleted mimetibody, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetibody of the invention.

XX
SQ Sequence 12 AA;

Query Match 95.8%; Score 46; DB 8; Length 12;
Best Local Similarity 88.9%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
| | | | | | |
Db 4 RPLPPLPPP 12

RESULT 44
AAR93386
ID AAR93386 standard; peptide; 14 AA.
XX
AC AAR93386;
XX
DT 30-APR-1996 (first entry)
XX
DE GST-SRC SH3 protein tyrosine kinase derived peptide #6.
XX
KW SH3 ligand; SH3 binding agent; biased phage library;
KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
KW p67; complex; chronic myelogenous leukaemia; cancer.

XX
OS Synthetic.
XX
PN WO9524419-A1.
XX
PD 14-SEP-1995.
XX
PF 13-MAR-1995; 95WO-US003208.
XX
PR 11-MAR-1994; 94US-00209835.
PR 06-JAN-1995; 95US-00369832.
XX
PA (ARIA-) ARIAD PHARM INC.
XX
PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
XX
DR WPI; 1995-328231/42.
XX
PT Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.

XX
PS Disclosure; Fig 4; 74pp; English.
XX
CC The sequences given in AAR93381-443 represent peptides which are SH3
CC ligands/SH3 binding agents. They represent a biased phage library which
CC comprises five random amino acids flanking the nonapeptide -RPLPPLPPP

CC which was identified as a recognition sequence for the src SH3 domain.
CC These sequences were identified using the method of the invention. The
CC method comprises contacting the SH3 domain with a mixture of peptides
CC under conditions permitting a ligand to bind to an SH3 domain to form a
CC complex. Any unbound peptides are removed and the complexed peptide
CC ligands are dissociated from the complexes. The selected peptides are
CC enriched by re-contacting them with the SH3 domain and then candidates
CC which bind to the SH3 domain are detected. The isolated SH3 binding
CC peptides may be used in the diagnosis, prevention and treatment of
CC conditions or diseases resulting from cellular processes mediated by an
CC SH3-based interaction. Such diseases include Paget's disease. Other
CC conditions treatable with these peptides include restenosis, rheumatoid
CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
CC p47 and p67 complex is implicated, etc
XX
SQ Sequence 14 AA;

Query Match 95.8%; Score 46; DB 2; Length 14;
Best Local Similarity 88.9%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RPLPPLPXP 12
Db 6 RPLPPLPPP 14
| | | | | | | |

RESULT 45
AAR93404
ID AAR93404 standard; peptide; 14 AA.
XX
AC AAR93404;
XX
DT 30-APR-1996 (first entry)
XX
DE GST-FYN SH3 protein tyrosine kinase derived peptide #14.
XX
KW SH3 ligand; SH3 binding agent; biased phage library;
KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
KW p67; complex; chronic myelogenous leukaemia; cancer.
XX
OS Synthetic.
XX
PN WO9524419-A1.
XX
PD 14-SEP-1995.
XX
PF 13-MAR-1995; 95WO-US003208.
XX
PR 11-MAR-1994; 94US-00209835.
PR 06-JAN-1995; 95US-00369832.
XX
PA (ARIA-) ARIAD PHARM INC.
XX
PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
XX
DR WPI; 1995-328231/42.
XX
PT Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.
XX
PS Disclosure; Fig 4; 74pp; English.
XX
CC The sequences given in AAR93381-443 represent peptides which are SH3
CC ligands/SH3 binding agents. They represent a biased phage library which
CC comprises five random amino acids flanking the nonapeptide -RPLPPLPPP
CC which was identified as a recognition sequence for the src SH3 domain.
CC These sequences were identified using the method of the invention. The
CC method comprises contacting the SH3 domain with a mixture of peptides
CC under conditions permitting a ligand to bind to an SH3 domain to form a
CC complex. Any unbound peptides are removed and the complexed peptide
CC ligands are dissociated from the complexes. The selected peptides are

CC enriched by re-contacting them with the SH3 domain and then candidates
CC which bind to the SH3 domain are detected. The isolated SH3 binding
CC peptides may be used in the diagnosis, prevention and treatment of
CC conditions or diseases resulting from cellular processes mediated by an
CC SH3-based interaction. Such diseases include Paget's disease. Other
CC conditions treatable with these peptides include restenosis, rheumatoid
CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
CC p47 and p67 complex is implicated, etc
XX
SQ Sequence 14 AA;

Query Match 95.8%; Score 46; DB 2; Length 14;
Best Local Similarity 88.9%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RPLPPLPXP 12
Db 6 RPLPPLPPP 14
| | | | | | | |

RESULT 46
AAR93450
ID AAR93450 standard; peptide; 14 AA.
XX
AC AAR93450;
XX
DT 30-APR-1996 (first entry)
XX
DE GST-PI3K SH3 protein tyrosine kinase derived peptide #7.
XX
KW SH3 ligand; SH3 binding agent; biased phage library;
KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
KW p67; complex; chronic myelogenous leukaemia; cancer.
XX
OS Synthetic.
XX
PN WO9524419-A1.
XX
PD 14-SEP-1995.
XX
PF 13-MAR-1995; 95WO-US003208.
XX
PR 11-MAR-1994; 94US-00209835.
PR 06-JAN-1995; 95US-00369832.
XX
PA (ARIA-) ARIAD PHARM INC.
XX
PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
XX
DR WPI; 1995-328231/42.
XX
PT Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.
XX
PS Disclosure; Fig 4; 74pp; English.
XX
CC The sequences given in AAR93444-56 represent peptides which are SH3
CC ligands/SH3 binding agents. They represent a biased phage library which
CC comprises five random amino acids flanking the nonapeptide -RPLPPLPPP
CC which was identified as a recognition sequence for the src SH3 domain.
CC These sequences were identified using the method of the invention. The
CC method comprises contacting the SH3 domain with a mixture of peptides
CC under conditions permitting a ligand to bind to an SH3 domain to form a
CC complex. Any unbound peptides are removed and the complexed peptide
CC ligands are dissociated from the complexes. The selected peptides are
CC enriched by re-contacting them with the SH3 domain and then candidates
CC which bind to the SH3 domain are detected. The isolated SH3 binding
CC peptides may be used in the diagnosis, prevention and treatment of
CC conditions or diseases resulting from cellular processes mediated by an
CC SH3-based interaction. Such diseases include Paget's disease. Other
CC conditons treatable with these peptides include restenosis, rheumatoid

CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
CC p47 and p67 complex is implicated, etc
XX
SQ Sequence 14 AA;
Query Match 95.8%; Score 46; DB 2; Length 14;
Best Local Similarity 88.9%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 RPLPPLPXP 12
Db 6 RPLPPLPPP 14
RESULT 47
AAR93388
ID AAR93388 standard; peptide; 14 AA.
XX
AC AAR93388;
XX
DT 30-APR-1996 (first entry)
DE GST-SRC SH3 protein tyrosine kinase derived peptide #8.
XX
KW SH3 ligand; SH3 binding agent; biased phage library;
KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
KW p67; complex; chronic myelogenous leukaemia; cancer.
XX
OS Synthetic.
XX
PN WO9524419-A1.
XX
PD 14-SEP-1995.
XX
PF 13-MAR-1995; 95WO-US003208.
XX
PR 11-MAR-1994; 94US-00209835.
PR 06-JAN-1995; 95US-00369832.
XX
PA (ARIA-) ARIAD PHARM INC.
XX
PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
XX WPI; 1995-328231/42.
XX
PT Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.
XX
PS Disclosure; Fig 4; 74pp; English.
XX
CC The sequences given in AAR93381-443 represent peptides which are SH3
CC ligands/SH3 binding agents. They represent a biased phage library which
CC comprises five random amino acids flanking the nonapeptide -RPLPPLPPP
CC which was identified as a recognition sequence for the src SH3 domain.
CC These sequences were identified using the method of the invention. The
CC method comprises contacting the SH3 domain with a mixture of peptides
CC under conditions permitting a ligand to bind to an SH3 domain to form a
CC complex. Any unbound peptides are removed and the complexed peptide
CC ligands are dissociated from the complexes. The selected peptides are
CC enriched by re-contacting them with the SH3 domain and then candidates
CC which bind to the SH3 domain are detected. The isolated SH3 binding
CC peptides may be used in the diagnosis, prevention and treatment of
CC conditions or diseases resulting from cellular processes mediated by an
CC SH3-based interaction. Such diseases include Paget's disease. Other
CC conditions treatable with these peptides include restenosis, rheumatoid
CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
CC p47 and p67 complex is implicated, etc
XX
SQ Sequence 14 AA;
Query Match 95.8%; Score 46; DB 2; Length 14;
Best Local Similarity 88.9%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Best Local Similarity 88.9%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 RPLPPLPXP 12
Db 6 RPLPPLPPP 14
RESULT 48
AAR93392
ID AAR93392 standard; peptide; 14 AA.
XX
AC AAR93392;
XX
DT 30-APR-1996 (first entry)
DE GST-FYN SH3 protein tyrosine kinase derived peptide #2.
XX
KW SH3 ligand; SH3 binding agent; biased phage library;
KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
KW p67; complex; chronic myelogenous leukaemia; cancer.
XX
OS Synthetic.
XX
PN WO9524419-A1.
XX
PD 14-SEP-1995.
XX
PF 13-MAR-1995; 95WO-US003208.
XX
PR 11-MAR-1994; 94US-00209835.
PR 06-JAN-1995; 95US-00369832.
XX
PA (ARIA-) ARIAD PHARM INC.
XX
PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
XX WPI; 1995-328231/42.
XX
PT Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.
XX
PS Disclosure; Fig 4; 74pp; English.
XX
CC The sequences given in AAR93381-443 represent peptides which are SH3
CC ligands/SH3 binding agents. They represent a biased phage library which
CC comprises five random amino acids flanking the nonapeptide -RPLPPLPPP
CC which was identified as a recognition sequence for the src SH3 domain.
CC These sequences were identified using the method of the invention. The
CC method comprises contacting the SH3 domain with a mixture of peptides
CC under conditions permitting a ligand to bind to an SH3 domain to form a
CC complex. Any unbound peptides are removed and the complexed peptide
CC ligands are dissociated from the complexes. The selected peptides are
CC enriched by re-contacting them with the SH3 domain and then candidates
CC which bind to the SH3 domain are detected. The isolated SH3 binding
CC peptides may be used in the diagnosis, prevention and treatment of
CC conditions or diseases resulting from cellular processes mediated by an
CC SH3-based interaction. Such diseases include Paget's disease. Other
CC conditions treatable with these peptides include restenosis, rheumatoid
CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
CC p47 and p67 complex is implicated, etc
XX
SQ Sequence 14 AA;
Query Match 95.8%; Score 46; DB 2; Length 14;
Best Local Similarity 88.9%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 RPLPPLPXP 12
Db 6 RPLPPLPPP 14


```
RESULT 49
AAR93400
ID AAR93400 standard; peptide; 14 AA.
XX AC
XX AAR93400;
XX AC
XX SH3 ligand; SH3 binding agent; biased phage library;
DT 30-APR-1996 (first entry)
XX GST-FYN SH3 protein tyrosine kinase derived peptide #10.
DE
XX
XX SH3 ligand; SH3 binding agent; biased phage library;
KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
KW p67; complex; chronic myelogenous leukaemia; cancer.
XX
XX Synthetic.
OS
XX WO9524419-A1.
PN
XX 14-SEP-1995.
PD
XX 13-MAR-1995; 95WO-US003208.
PF
XX 11-MAR-1994; 94US-00209835.
PR
PR 06-JAN-1995; 95US-00369832.
XX
XX (ARIA-) ARIAD PHARM INC.
PA
XX Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
PI
XX WPI; 1995-328231/42.
DR
XX Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.
XX
XX Disclosure; Fig 4; 74pp; English.
PS
XX The sequences given in AAR93381-443 represent peptides which are SH3
CC ligands/SH3 binding agents. They represent a biased phage library which
CC comprises five random amino acids flanking the nonapeptide -RPLPPLPPP
CC which was identified as a recognition sequence for the src SH3 domain.
CC These sequences were identified using the method of the invention. The
CC method comprises contacting the SH3 domain with a mixture of peptides
CC under conditions permitting a ligand to bind to an SH3 domain to form a
CC complex. Any unbound peptides are removed and the complexed peptide
CC ligands are dissociated from the complexes. The selected peptides are
CC enriched by re-contacting them with the SH3 domain and then candidates
CC which bind to the SH3 domain are detected. The isolated SH3 binding
CC peptides may be used in the diagnosis, prevention and treatment of
CC conditions or diseases resulting from cellular processes mediated by an
CC SH3-based interaction. Such diseases include Paget's disease. Other
CC conditions treatable with these peptides include restenosis, rheumatoid
CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
CC p47 and p67 complex is implicated, etc
XX
SQ Sequence 14 AA;
Query Match 95.8%; Score 46; DB 2; Length 14;
Best Local Similarity 88.9%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 RPLPPLPPX 12
Db 6 RPLPPLPPP 14
Search completed: April 6, 2006, 09:30:42
Job time : 168.211 secs
```

```
AC AAR93428;
XX
DT 30-APR-1996 (first entry)
XX
DE GST-YES SH3 protein tyrosine kinase derived peptide #4.
XX
XX SH3 ligand; SH3 binding agent; biased phage library;
KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
KW p67; complex; chronic myelogenous leukaemia; cancer.
XX
XX Synthetic.
OS
XX WO9524419-A1.
PN
XX 14-SEP-1995.
PD
XX 13-MAR-1995; 95WO-US003208.
PF
XX 11-MAR-1994; 94US-00209835.
PR
PR 06-JAN-1995; 95US-00369832.
XX
XX (ARIA-) ARIAD PHARM INC.
PA
XX Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
PI
XX WPI; 1995-328231/42.
DR
XX Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.
XX
XX Disclosure; Fig 4; 74pp; English.
PS
XX The sequences given in AAR93381-443 represent peptides which are SH3
CC ligands/SH3 binding agents. They represent a biased phage library which
CC comprises five random amino acids flanking the nonapeptide -RPLPPLPPP
CC which was identified as a recognition sequence for the src SH3 domain.
CC These sequences were identified using the method of the invention. The
CC method comprises contacting the SH3 domain with a mixture of peptides
CC under conditions permitting a ligand to bind to an SH3 domain to form a
CC complex. Any unbound peptides are removed and the complexed peptide
CC ligands are dissociated from the complexes. The selected peptides are
CC enriched by re-contacting them with the SH3 domain and then candidates
CC which bind to the SH3 domain are detected. The isolated SH3 binding
CC peptides may be used in the diagnosis, prevention and treatment of
CC conditions or diseases resulting from cellular processes mediated by an
CC SH3-based interaction. Such diseases include Paget's disease. Other
CC conditions treatable with these peptides include restenosis, rheumatoid
CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
CC p47 and p67 complex is implicated, etc
XX
SQ Sequence 14 AA;
Query Match 95.8%; Score 46; DB 2; Length 14;
Best Local Similarity 88.9%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 RPLPPLPPX 12
Db 6 RPLPPLPPP 14
Search completed: April 6, 2006, 09:30:42
Job time : 168.211 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.
OM protein - protein search, using sw model
Run on: April 6, 2006, 09:35:13 ; Search time 21.7895 Seconds
(without alignments)
52.989 Million cell updates/sec

Title: US-10-632-388-307
Perfect score: 48
Sequence: 1 XXXRPLPLPPXP 12
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : PIR 80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result NO.	Score	Query Match	Length DB ID	Description
1	47	97.9	613 2 A56031	potassium channel
2	43	89.6	162 2 T07173	hypothetical prote
3	42	87.5	198 2 E89008	protein W08A12.3 {
4	42	87.5	281 2 I38707	Fas ligand - human
5	42	87.5	473 2 S50755	hypothetical prote
6	41.	85.4	339 2 JC7509	glycoprotein VI-1
7	41	85.4	348 2 D88088	protein B0454.1 {i
8	41	85.4	459 2 S03116	gene 33 protein, h
9	41	85.4	477 2 T46917	hypothetical prote
10	41	85.4	535 2 A46101	protein-tyrosine-p
11	41	85.4	548 2 B46101	protein-tyrosine-p
12	41	85.4	627 2 T26064	hypothetical prote
13	41	85.4	663 1 TVMVR	protein-tyrosine k
14	41	85.4	894 2 F84870	hypothetical prote
15	41	85.4	1225 2 T16346	hypothetical prote
16	40	83.3	134 2 T36365	proline-rich prote
17	40	83.3	156 1 GNVQL2	genome-linked prot
18	40	83.3	156 1 GNVQLL	genome-linked prot
19	40	83.3	156 1 GNVQWA	genome-linked prot
20	40	83.3	215 2 T15811	hypothetical prote
21	40	83.3	224 2 T16669	hypothetical prote
22	40	83.3	245 2 T29543	hypothetical prote
23	40	83.3	273 2 C70551	hypothetical prote
24	40	83.3	574 2 T41395	probable dna polym
25	40	83.3	775 2 C81594	hypothetical prote
26	40	83.3	775 2 D86549	hypothetical prote
27	40	83.3	775 2 B72074	hypothetical prote
28	40	83.3	793 1 S60735	splicing factor SF
29	39	81.2	93 2 G30010	hypothetical ORF-8

30	81.2	153	2	F83532	conserved hypothet
31	81.2	173	2	T19341	hypothetical prote
32	81.2	373	2	A70856	probable lppz prot
33	81.2	409	2	S60975	hypothetical prote
34	81.2	446	2	A34794	B-cell CLL/lymphom
35	81.2	472	2	A75464	probable zinc meta
36	81.2	514	2	C49507	potassium channel
37	81.2	602	2	A49507	potassium channel
38	81.2	832	2	T49494	condensin complex
39	81.2	1106	2	T31742	hypothetical prote
40	81.2	1211	2	T42230	AF4 protein - mous
41	81.2	1217	2	T42625	AF-4 protein - mou
42	81.2	1733	1	B45344	probable nuclear a
43	81.2	1958	2	B40505	hypothetical prote
44	81.2	2493	2	S72349	nonstructural poly
45	81.2	2493	2	S26372	nonstructural poly
46	79.2	191	2	E87601	OmpA family protei
47	79.2	217	2	T04353	DNA binding protei
48	79.2	222	2	C75539	conserved hypothet
49	79.2	308	2	A38582	pollen allergen pr
50	79.2	318	1	A31978	ADP,ATP carrier pr
51	79.2	336	2	E47301	Vir10 homolog - Bo
52	79.2	407	2	H87341	OmpA family protei
53	79.2	408	1	G65132	hypothetical 44.6
54	79.2	443	2	I49140	p62 ras-GAP associ
55	79.2	527	2	I84483	protein-tyrosine k
56	79.2	527	2	I49133	protein-tyrosine k
57	79.2	668	2	S56909	polymyxin B resist
58	79.2	925	2	S50490	hypothetical prote
59	79.2	1006	2	T42762	probable alpha-man
60	79.2	1386	2	T49316	profilaggrin relat
61	79.2	1560	2	T00080	hypothetical prote
62	77.1	68	2	PW0050	synaptotjanin 2 eps
63	77.1	79	2	T44145	B2 protein [import
64	77.1	101	2	T46506	hypothetical prote
65	77.1	104	2	D72542	hypothetical prote
66	77.1	109	2	G86433	protein Ti7H7.7 {i
67	77.1	124	2	G95982	hypothetical expor
68	77.1	161	2	T16207	hypothetical prote
69	77.1	166	2	T36616	hypothetical prote
70	77.1	175	2	S65828	probable movement
71	77.1	298	2	A75460	conserved hypothet
72	77.1	408	2	T32767	hypothetical prote
73	77.1	454	2	A75444	hypothetical prote
74	77.1	474	2	T20108	hypothetical prote
75	77.1	487	2	S42442	nuclear protein EB
76	77.1	547	2	C96828	unknown protein F1
77	77.1	667	2	T17221	hypothetical prote
78	77.1	695	2	T24950	hypothetical prote
79	77.1	745	2	T51370	hypothetical prote
80	77.1	876	2	T49801	hypothetical prote
81	77.1	1140	2	D88690	hypothetical prote
82	77.1	1201	2	A57369	protein F41H10.3 {
83	77.1	1262	2	T13353	anillin - fruit fl
84	77.1	1464	2	T13716	protein stn-B - fr
85	77.1	1819	2	T32008	bazooka gene prote
86	76.0	226	2	I39594	hypothetical prote
87	76.0	882	2	I38912	exEB protein - Aer
88	76.0	890	1	A53743	receptor tyrosine
89	75.0	104	1	R5PM25	protein-tyrosine k
90	75.0	121	2	T04536	ribosomal protein
91	75.0	123	2	AH2707	hypothetical prote
92	75.0	169	2	H72470	conserved hypothet
93	75.0	196	2	T00702	hypothetical prote
94	75.0	199	2	T40919	hypothetical prote
95	75.0	217	2	S09754	probable signal re
96	75.0	224	2	D72861	hypothetical prote
97	75.0	231	2	F82577	AcOrf-91 protein -
98	75.0	235	2	G83625	TonB protein XF228
99	75.0	238	2	T32889	hypothetical prote
100	75.0	256	2	A35340	hypothetical prote
101	75.0	265	2	H75560	H+-transporting tw
102	75.0	267	2	JQ1752	oxidoreductase, sh
					hypothetical 30.6K

103 36 75.0 271 2 D70915 hypothetical prote
104 36 75.0 276 2 A12292 hypothetical prote
105 36 75.0 277 2 I38857 microtubule-associ
106 36 75.0 285 2 T34615 NADH2 dehydrogenas
107 36 75.0 291 2 G84494 hypothetical prote
108 36 75.0 319 2 B87696 hypothetical prote
109 36 75.0 341 2 E69463 type I restriction
110 36 75.0 346 2 S76923 hypothetical prote
111 36 75.0 352 2 T18794 hypothetical prote
112 36 75.0 361 2 S19552 potassium channel
113 36 75.0 378 2 T28112 hypothetical prote
114 36 75.0 380 2 T09661 ascorbate oxidase
115 36 75.0 393 2 T32533 hypothetical prote
116 36 75.0 417 2 G64417 hypothetical prote
117 36 75.0 426 2 F95058 hypothetical prote
118 36 75.0 428 2 S76184 hypothetical prote
119 36 75.0 431 2 S20065 nuclear factor I-X
120 36 75.0 435 2 A42672 choline kinase (EC
121 36 75.0 441 2 S37881 hypothetical prote
122 36 75.0 443 2 A38219 GAP-associated tyr
123 36 75.0 453 2 JX0342 choline kinase (EC
124 36 75.0 485 2 A33647 sulfated surface g
125 36 75.0 522 2 H97927 type 1 site-specif
126 36 75.0 524 2 A75588 probable protein k
127 36 75.0 577 2 T09024 proline-rich prote
128 36 75.0 598 2 T02265 hypothetical prote
129 36 75.0 604 2 H81110 sulfite reductase
130 36 75.0 606 1 UZAD12 terminal protein p
131 36 75.0 609 2 T28736 hypothetical prote
132 36 75.0 618 2 T05518 hypothetical prote
133 36 75.0 645 2 T16078 hypothetical prote
134 36 75.0 649 2 AE1033 hypothetical prote
135 36 75.0 653 1 UZADP2 terminal protein p
136 36 75.0 653 1 UZADP5 terminal protein p
137 36 75.0 679 2 A42073 potassium channel
138 36 75.0 679 2 B75262 conserved hypothet
139 36 75.0 691 2 F91251 probable tape meas
140 36 75.0 715 1 TNBE77 77K alpha trans-in
141 36 75.0 745 2 T39952 hypothetical prote
142 36 75.0 745 2 I56546 Shaw type potassiu
143 36 75.0 772 2 T13078 KIAA0992 protein -
144 36 75.0 818 2 T02231 probable isoamylas
145 36 75.0 822 2 T09221 exocyst complex pr
146 36 75.0 833 2 T01547 probable phospholi
147 36 75.0 885 2 S67660 hypothetical prote
148 36 75.0 927 2 T24031 hypothetical prote
149 36 75.0 988 2 F86316 protein T10022.13
150 36 75.0 992 1 GNWVR3 structural polypro
151 36 75.0 1013 2 JNC2314 chitin synthase (E
152 36 75.0 1014 2 JE0333 klotho protein - r
153 36 75.0 1048 2 T31425 C-terminal domain-
154 36 75.0 1102 2 JC6316 probable protein k
155 36 75.0 1173 2 T31421 C-terminal domain-
156 36 75.0 1200 2 T48194 hypothetical prote
157 36 75.0 1239 2 G71266 probable ATP-depen
158 36 75.0 1262 2 T25168 hypothetical prote
159 36 75.0 1264 2 A36858 G2R protein - vari
160 36 75.0 1334 2 T50568 probable multi-dom
161 36 75.0 1392 2 T51947 probable transcrip
162 36 75.0 1400 2 T52359 hypothetical prote
163 36 75.0 1422 2 T24212 hypothetical prote
164 36 75.0 1445 2 A59437 KIAA1204 protein l
165 36 75.0 1466 2 A36426 SPA2 protein - yea
166 36 75.0 1557 2 T13160 protein CNK - frui
167 36 75.0 1575 2 T18545 lysobactin synthet
168 36 75.0 1742 2 T49451 kinesin-like prote
169 36 75.0 1874 1 JQ0533 genome polyprotein
170 36 75.0 1896 2 B72175 D15R protein - var
171 36 75.0 1897 2 T28621 hypothetical prote
172 36 75.0 2105 2 T18968 probable serine-ty
173 36 75.0 2133 2 T30637 hypothetical prote
174 36 75.0 2459 2 AF2136 peptide synthetase
175 36 75.0 2774 2 A43359 microtubule-associ

176 36 75.0 3175 1 RRWVEV genome polyprotein
177 35 72.9 74 2 F36800 hypothetical prote
178 35 72.9 79 2 AE2788 hypothetical prote
179 35 72.9 117 2 F72669 hypothetical prote
180 35 72.9 134 2 A72730 hypothetical prote
181 35 72.9 150 2 E72587 hypothetical prote
182 35 72.9 152 2 D95961 hypothetical prote
183 35 72.9 152 2 C72662 ACWNPV orf91 - Bom
184 35 72.9 154 2 T41831 hypothetical prote
185 35 72.9 154 2 E72624 hypothetical prote
186 35 72.9 179 2 F83305 conserved hypothet
187 35 72.9 195 2 H75455 conserved hypothet
188 35 72.9 197 2 T33525 hypothetical prote
189 35 72.9 198 2 T35752 hypothetical prote
190 35 72.9 201 2 D70864 hypothetical prote
191 35 72.9 203 2 T47571 hypothetical prote
192 35 72.9 209 2 A48232 cysteine-rich exte
193 35 72.9 216 2 B37332 transforming prote
194 35 72.9 221 2 T15845 transforming prote
195 35 72.9 232 2 S24390 transforming prote
196 35 72.9 233 2 A37332 transforming prote
197 35 72.9 239 2 T36995 hypothetical prote
198 35 72.9 255 2 AH0321 conserved hypothet
199 35 72.9 257 2 S32101 PHLPSA protein - c
200 35 72.9 262 2 B72037 ct598 hypothetical
201 35 72.9 262 2 E86588 CT598 hypothetical
202 35 72.9 262 2 A81504 conserved hypothet
203 35 72.9 264 2 PQ0478 pistil extensin-li
204 35 72.9 265 1 B45393 polypeptide VI pre
205 35 72.9 268 2 T30630 hypothetical prote
206 35 72.9 269 2 T30468 hypothetical prote
207 35 72.9 279 2 D96792 probable RNA-bind
208 35 72.9 293 2 C83845 superoxide dismuta
209 35 72.9 309 2 G87498 hypothetical prote
210 35 72.9 313 1 FOVDA gag polyprotein -
211 35 72.9 316 2 T31880 hypothetical prote
212 35 72.9 321 2 F69451 cofactor modifying
213 35 72.9 358 2 A75561 conserved hypothet
214 35 72.9 368 2 T06093 conserved hypothet
215 35 72.9 369 2 B70968 hypothetical prote
216 35 72.9 380 2 T32944 hypothetical prote
217 35 72.9 384 2 H88924 protein F33E11.2 l
218 35 72.9 387 2 T21370 hypothetical prote
219 35 72.9 389 2 S27200 proline-rich prote
220 35 72.9 392 2 JC2206 sensory kinase - S
221 35 72.9 393 2 PQ0479 pistil extensin-li
222 35 72.9 396 2 T04100 glucose-6-phosphat
223 35 72.9 406 2 T44819 dolichyl-phosphate
224 35 72.9 415 1 A34170 acrosin (EC 3.4.21
225 35 72.9 421 1 S11674 acrosin (EC 3.4.21
226 35 72.9 426 2 JQ1696 pistil extensin-li
227 35 72.9 436 2 T01652 zinc finger protei
228 35 72.9 439 2 S51939 chitinase (EC 3.2.
229 35 72.9 440 2 T41766 ARIF-1 orf20/21 -
230 35 72.9 445 2 H96560 hypothetical prote
231 35 72.9 451 2 T30603 perlecan homolog 2
232 35 72.9 456 2 S23104 choline kinase - h
233 35 72.9 457 2 T50233 probable DNA-direc
234 35 72.9 471 2 E83592 hypothetical prote
235 35 72.9 490 2 A35312 potassium channel
236 35 72.9 523 2 A31948 carnitine octanoyl
237 35 72.9 530 2 A45690 transactivator EBN
238 35 72.9 530 2 G70904 hypothetical prote
239 35 72.9 543 2 D83262 hypothetical prote
240 35 72.9 550 2 T36746 probable serine/th
241 35 72.9 563 2 A87378 hypothetical prote
242 35 72.9 586 2 T29695 hypothetical prote
243 35 72.9 589 2 T05074 hypothetical prote
244 35 72.9 598 2 S66669 potassium channel
245 35 72.9 599 2 T10798 pterophorin-S - Vo
246 35 72.9 602 2 JH0166 potassium voltage-
247 35 72.9 602 2 F84432 probable C2H2-type
248 35 72.9 612 2 S60025 carnitine octanoyl

249	35	72.9	641	2	A54407	alpha-mannosidase	322	34	70.8	226	2	S41032	hypothetical prote
250	35	72.9	645	2	G01205	TYL protein - huma	323	34	70.8	226	2	C45343	hypothetical prote
251	35	72.9	650	2	B87791	protein B0207.1 [i	324	34	70.8	226	2	T42558	tegument protein 1
252	35	72.9	664	2	T01368	hypothetical prote	325	34	70.8	229	2	F84370	hypothetical prote
253	35	72.9	678	2	T04832	probable serine/th	326	34	70.8	238	2	T05263	hypothetical prote
254	35	72.9	681	2	JC5929	serine/arginine-ri	327	34	70.8	241	2	T22216	hypothetical prote
255	35	72.9	706	2	A45990	junctional sarcopl	328	34	70.8	245	2	T20831	hypothetical prote
256	35	72.9	708	2	D96711	hypothetical prote	329	34	70.8	247	2	G75356	hypothetical membr
257	35	72.9	734	2	T04876	hypothetical prote	330	34	70.8	248	2	F83558	hypothetical prote
258	35	72.9	749	2	S61643	probable membrane	331	34	70.8	251	2	F83101	conserved hypothet
259	35	72.9	774	1	JQ0550	1,4-alpha-glucan b	332	34	70.8	257	2	C72651	hypothetical prote
260	35	72.9	779	2	AC2249	heterocyst differe	333	34	70.8	275	2	G95962	probable taurine u
261	35	72.9	798	2	T21369	hypothetical prote	334	34	70.8	282	2	T26635	hypothetical prote
262	35	72.9	823	2	A36378	probable transcrip	335	34	70.8	283	2	S13383	hydroxyproline-ric
263	35	72.9	847	2	F96531	hypothetical prote	336	34	70.8	289	2	C96752	probable DNA-bindi
264	35	72.9	868	2	T20239	hypothetical prote	337	34	70.8	297	2	G70708	probable purC prot
265	35	72.9	899	2	T35299	probable pyruvate	338	34	70.8	298	2	H87533	peptidase, M23/M37
266	35	72.9	915	2	T12526	hypothetical prote	339	34	70.8	301	2	JQ1663	hybrid proline-ric
267	35	72.9	939	2	S28394	probable serine/th	340	34	70.8	301	2	T18788	hypothetical prote
268	35	72.9	944	2	T47246	chitin synthase (E	341	34	70.8	308	2	AH2896	transcription regu
269	35	72.9	974	2	E59434	Rho GTPase activat	342	34	70.8	308	2	B97672	rok family protein
270	35	72.9	1013	2	T33470	hypothetical prote	343	34	70.8	311	2	H70911	hypothetical prote
271	35	72.9	1041	2	T31097	chitin synthase (E	344	34	70.8	312	2	S04281	psbA intron 2 prot
272	35	72.9	1043	2	T19734	hypothetical prote	345	34	70.8	315	2	T06806	proline rich prote
273	35	72.9	1050	2	T27753	hypothetical prote	346	34	70.8	322	2	T23891	hypothetical prote
274	35	72.9	1175	2	S51005	protein-tyrosine-p	347	34	70.8	330	2	T05717	probable extensin
275	35	72.9	1176	2	I58345	protein tyrosine p	348	34	70.8	335	2	S64352	hypothetical prote
276	35	72.9	1183	2	S63046	probable membrane	349	34	70.8	340	2	T36875	probable partition
277	35	72.9	1187	1	JC4155	protein-tyrosine-p	350	34	70.8	348	2	AD2247	hypothetical prote
278	35	72.9	1189	1	JC2366	protein-tyrosine-p	351	34	70.8	348	2	T47494	hypothetical prote
279	35	72.9	1245	1	VHVB	structural polypro	352	34	70.8	353	2	B36963	hypothetical prote
280	35	72.9	1245	1	VHVB2	structural polypro	353	34	70.8	358	2	B86141	bcsA 5'-region pro
281	35	72.9	1249	2	T14150	vesicle associated	354	34	70.8	359	2	T13478	protein T25K16.3 [
282	35	72.9	1258	2	JC5765	inositol polyphosp	355	34	70.8	361	2	T26610	hypothetical prote
283	35	72.9	1353	2	T00249	ich1 protein - ink	356	34	70.8	369	2	AG1950	hypothetical prote
284	35	72.9	1428	2	T08852	lustrin A - Califo	357	34	70.8	372	1	T04157	dihydrokaempferol
285	35	72.9	1440	2	T27942	lin-15B protein -	358	34	70.8	380	2	T00683	hypothetical prote
286	35	72.9	1602	2	T31671	Rab3 GDP/GTP excha	359	34	70.8	386	2	E72645	hypothetical prote
287	35	72.9	1657	2	T19536	hypothetical prote	360	34	70.8	396	1	A58938	surface protein rh
288	35	72.9	1662	1	H71402	probable kinesin -	361	34	70.8	396	2	S41224	hnRNP protein - Af
289	35	72.9	1870	2	S37671	MHC class III hist	362	34	70.8	398	2	T37083	probable transposa
290	35	72.9	1872	2	S36152	MHC class III hist	363	34	70.8	407	2	S77922	shaggy-like protei
291	35	72.9	1879	2	T15962	hypothetical prote	364	34	70.8	409	2	D42971	terpredoxin reduct
292	35	72.9	2186	2	T13169	tiggrin - fruit fl	365	34	70.8	409	2	S51105	shaggy protein kin
293	35	72.9	2241	2	S09811	hypothetical prote	366	34	70.8	410	2	C70889	hypothetical prote
294	35	72.9	2453	2	S60254	nuclear receptor c	367	34	70.8	412	2	S71266	shaggy-like protei
295	35	72.9	3119	2	I49729	HD protein - mouse	368	34	70.8	412	2	A84715	probable shaggy-li
296	34.5	71.9	59	2	A36589	bactenecin 7 - bov	369	34	70.8	418	2	F82638	hypothetical prote
297	34	70.8	81	2	B70727	hypothetical prote	370	34	70.8	420	2	A96613	probable glycogen
298	34	70.8	101	2	S29093	SMD3 protein - yea	371	34	70.8	420	2	T36193	probable salicylat
299	34	70.8	101	2	H72685	hypothetical prote	372	34	70.8	421	2	S51938	protein kinase AtK
300	34	70.8	106	2	T47814	hypothetical prote	373	34	70.8	423	2	T24383	hypothetical prote
301	34	70.8	111	2	T00688	hypothetical prote	374	34	70.8	432	2	AB2222	twitching motility
302	34	70.8	117	2	AF3319	hypothetical prote	375	34	70.8	435	2	T46443	hypothetical prote
303	34	70.8	130	2	E72697	hypothetical prote	376	34	70.8	447	2	F86232	hypothetical prote
304	34	70.8	135	2	PC4252	hypothetical 135 p	377	34	70.8	447	2	T20249	hypothetical prote
305	34	70.8	141	2	JQ1669	hypothetical 16.1K	378	34	70.8	451	2	AI3623	hypothetical prote
306	34	70.8	144	2	H72464	hypothetical prote	379	34	70.8	452	2	D75572	acriflavin resista
307	34	70.8	145	2	T48552	glutaredoxin-like	380	34	70.8	463	2	S41495	dC stretch-binding
308	34	70.8	150	2	B72705	hypothetical prote	381	34	70.8	464	2	A54143	kappa-B motif-bind
309	34	70.8	152	2	JQ1244	coat protein 2 - b	382	34	70.8	464	2	S43363	transformation upr
310	34	70.8	152	2	T00877	hypothetical prote	383	34	70.8	464	2	A47655	spliceosome-associ
311	34	70.8	160	2	D48232	cysteine-rich exte	384	34	70.8	469	1	T02254	shaggy protein kin
312	34	70.8	161	2	E48232	cysteine-rich exte	385	34	70.8	469	1	T02256	shaggy protein kin
313	34	70.8	165	2	C48232	cysteine-rich exte	386	34	70.8	471	1	T02297	shaggy protein kin
314	34	70.8	175	1	GNVQFL	genome-linked prot	387	34	70.8	471	2	T03601	shaggy protein kin
315	34	70.8	175	2	D70824	hypothetical prote	388	34	70.8	474	2	S57239	gene forked protei
316	34	70.8	190	2	T35570	hypothetical prote	389	34	70.8	476	2	S57963	methyl CpG binding
317	34	70.8	198	2	D70509	hypothetical prote	390	34	70.8	477	2	I38409	adenylyl cyclase-a
318	34	70.8	207	2	I53154	hypothetical prote	391	34	70.8	477	2	AE0608	conserved hypothet
319	34	70.8	211	2	B89716	scleraxis - mouse	392	34	70.8	478	2	F96594	unknown protein, 5
320	34	70.8	212	2	S74288	protein F45B8.3 [i	393	34	70.8	479	1	IXBE1F	alpha trans-induci
321	34	70.8	219	2	F87505	hypothetical prote	394	34	70.8	483	1	TVRTKA	protein kinase (EC

395 34 70.8 485 2 PC4427 Nck, Ash and phosph
396 34 70.8 490 1 IXBE33 alpha trans-induci
397 34 70.8 490 1 IXBE17 alpha trans-induci
398 34 70.8 490 2 JS0689 virion protein 16
399 34 70.8 491 2 T07598 proline-rich prote
400 34 70.8 491 2 AG1838 hypothetical prote
401 34 70.8 492 2 C96521 protein F21D18.18
402 34 70.8 498 2 AC0442 ABC transporter AT
403 34 70.8 499 2 S09880 hypothetical prote
404 34 70.8 502 2 A55197 Wiskott-Aldrich sy
405 34 70.8 508 2 T45867 hypothetical prote
406 34 70.8 520 1 QQCV6S hypothetical prote
407 34 70.8 529 2 T00677 hypothetical prote
408 34 70.8 534 2 A37483 F protein - Muraya
409 34 70.8 542 2 A82965 hypothetical prote
410 34 70.8 542 2 A44358 zyxin - chicken
411 34 70.8 555 2 T30349 structural protein
412 34 70.8 558 1 S68203 tyrosine N-monooxy
413 34 70.8 568 2 T40656 probable oxalyl-Co
414 34 70.8 579 2 E75275 hypothetical prote
415 34 70.8 586 2 H86914 conserved hypothet
416 34 70.8 587 2 T41653 probable transcrip
417 34 70.8 589 2 AC1446 probable portal pr
418 34 70.8 589 2 S60154 serine/threonine-s
419 34 70.8 589 2 S60159 serine/threonine-s
420 34 70.8 604 2 E81905 probable sulfite r
421 34 70.8 604 2 S39885 forked protein - f
422 34 70.8 606 2 S70829 primase - Myxococc
423 34 70.8 616 2 AE1973 serine/threonine k
424 34 70.8 621 2 B89451 probable zinc meta
425 34 70.8 626 1 NDBE61 exonuclease (EC 3.
426 34 70.8 632 2 S38042 probable purine nu
427 34 70.8 636 2 H95918 probable DNA ligas
428 34 70.8 637 2 T04552 hypothetical prote
429 34 70.8 641 2 S57236 forked protein 2.5
430 34 70.8 649 2 JN0809 drebrin E (clone g
431 34 70.8 651 2 T14763 hypothetical prote
432 34 70.8 654 2 T34960 cell division prot
433 34 70.8 658 2 T04831 probable serine/th
434 34 70.8 662 2 D40228 neurexin II-beta p
435 34 70.8 709 2 T46413 hypothetical prote
436 34 70.8 726 2 G83310 conserved hypothet
437 34 70.8 733 2 T12220 NADH2 dehydrogenas
438 34 70.8 736 2 I51691 dishevelled homolo
439 34 70.8 741 2 D83633 conserved hypothet
440 34 70.8 751 2 T34490 hypothetical prote
441 34 70.8 781 2 T02272 hypothetical prote
442 34 70.8 793 2 AH1094 probable peptidogl
443 34 70.8 810 2 S65226 probable membrane
444 34 70.8 814 2 T49207 receptor kinase-li
445 34 70.8 817 2 S51342 verprolin - yeast
446 34 70.8 827 2 T04789 hypothetical prote
447 34 70.8 839 2 T04859 extensin homolog F
448 34 70.8 847 2 S75975 hypothetical prote
449 34 70.8 848 2 E71404 hypothetical prote
450 34 70.8 869 2 T22422 hypothetical prote
451 34 70.8 872 2 T50369 probable serine/th
452 34 70.8 897 2 T02808 conserved hypothet
453 34 70.8 963 2 T48707 related to regulat
454 34 70.8 982 1 GNLJH2 pol polyprotein -
455 34 70.8 986 2 T33135 hypothetical prote
456 34 70.8 1038 2 S52522 hypothetical prote
457 34 70.8 1076 2 S50536 hypothetical prote
458 34 70.8 1091 2 S33596 protein-tyrosine k
459 34 70.8 1100 2 JC8033 leukocyte formin p
460 34 70.8 1111 2 T23047 hypothetical prote
461 34 70.8 1174 2 I38140 protein-tyrosine-p
462 34 70.8 1196 2 T09356 brassinosteroid-in
463 34 70.8 1215 2 S50428 probable Ca2+-tran
464 34 70.8 1344 2 T14316 rig-1 protein - mo
465 34 70.8 1360 2 T06699 zinc finger protei
466 34 70.8 1388 2 T00063 hypothetical prote
467 34 70.8 1425 2 E89303 protein C47E8.8 [i

468 34 70.8 1429 2 T13720 gene expanded prot
469 34 70.8 1436 2 S57238 forked protein 5.4
470 34 70.8 1449 2 S57237 forked protein 5.6
471 34 70.8 1459 2 T30196 kinesin motor prot
472 34 70.8 1675 2 T31473 hypothetical prote
473 34 70.8 1710 2 T14005 phospholipase D (E
474 34 70.8 1776 2 G86280 protein T5E21.13 [i
475 34 70.8 1872 2 T24683 hypothetical prote
476 34 70.8 2123 2 F86348 hypothetical prote
477 34 70.8 2212 2 T28157 erythrocyte membra
478 34 70.8 2649 2 T51023 hypothetical prote
479 34 70.8 7962 2 I38346 elastic titin - hu
480 33.5 69.8 1174 2 T49868 related to suppress
481 33 68.8 52 2 E98316 hypothetical prote
482 33 68.8 55 2 S29770 DNA-binding protei
483 33 68.8 69 2 D72852 actin rearrangemen
484 33 68.8 70 2 A96650 protein F24O1.6 [i
485 33 68.8 75 2 T26916 hypothetical prote
486 33 68.8 76 2 E97763 protein transport
487 33 68.8 89 1 NSBOH7 nonhistone chromos
488 33 68.8 89 2 S11349 nonhistone chromos
489 33 68.8 89 2 S33866 nonhistone chromos
490 33 68.8 90 1 NSCHH7 nonhistone chromos
491 33 68.8 90 2 S03700 nonhistone chromos
492 33 68.8 90 2 S01946 nonhistone chromos
493 33 68.8 105 1 NSCHH4 nonhistone chromos
494 33 68.8 115 2 T36886 hypothetical prote
495 33 68.8 125 2 A53692 synapsin I - mouse
496 33 68.8 126 2 JU0068 homeotic protein C
497 33 68.8 132 2 T49536 hypothetical prote
498 33 68.8 132 2 H96708 hypothetical prote
499 33 68.8 139 2 T05847 hypothetical prote
500 33 68.8 147 2 C75547 conserved hypothet
501 33 68.8 172 2 D41132 collagen-related p
502 33 68.8 175 2 D75514 hypothetical prote
503 33 68.8 187 2 C75558 acetyl-CoA carboxy
504 33 68.8 189 2 D49600 genome-linked prot
505 33 68.8 189 2 G70512 probable lppK prot
506 33 68.8 196 2 B48232 cysteine-rich exte
507 33 68.8 199 2 AB3502 phosphatidyl-N-met
508 33 68.8 199 2 AB2171 hypothetical prote
509 33 68.8 206 2 T09503 probable heme tran
510 33 68.8 207 2 C84631 nodulin-like prote
511 33 68.8 210 2 T50713 urease accessory p
512 33 68.8 212 2 E91046 transcription regu
513 33 68.8 212 2 F65023 gcvR protein - Esc
514 33 68.8 212 2 A85891 transcription regu
515 33 68.8 214 2 S00763 H+-transporting tw
516 33 68.8 215 2 A69436 LSU ribosomal prot
517 33 68.8 219 2 AH2953 conserved hypothet
518 33 68.8 221 2 T07176 extensin homolog -
519 33 68.8 230 2 T29591 hypothetical prote
520 33 68.8 240 2 T25814 hypothetical prote
521 33 68.8 248 2 T40149 hypothetical prote
522 33 68.8 254 2 E98329 hypothetical prote
523 33 68.8 256 2 JQ1144 H+-transporting tw
524 33 68.8 257 2 T29183 hypothetical prote
525 33 68.8 259 2 T22596 hypothetical prote
526 33 68.8 263 2 F70843 probable sdhB prot
527 33 68.8 264 2 T06789 hydroxyproline-ric
528 33 68.8 265 2 AG2855 conserved hypothet
529 33 68.8 265 2 D97632 hypothetical prote
530 33 68.8 267 2 AC3133 hypothetical prote
531 33 68.8 269 2 JC4899 proline rich prote
532 33 68.8 272 2 AB2988 conserved hypothet
533 33 68.8 284 2 AB1919 hypothetical prote
534 33 68.8 286 2 S34665 collagen, cuticula
535 33 68.8 288 2 T00473 hypothetical prote
536 33 68.8 293 2 S57376 hypothetical prote
537 33 68.8 293 2 AE2018 hypothetical prote
538 33 68.8 294 2 T34537 hypothetical prote
539 33 68.8 294 2 A55477 survival motor neu
540 33 68.8 294 2 A96155 hypothetical prote

541	33	68.8	295	2	C95285	probable LysR-type	614	33	68.8	625	1	A43030	protein-tyrosine k
542	33	68.8	296	1	WMADL2	late 100K protein	615	33	68.8	627	2	D75393	serine proteinase,
543	33	68.8	299	2	G82616	TonB protein XF195	616	33	68.8	631	2	T32761	hypothetical prote
544	33	68.8	314	2	F96527	protein F27J15.20	617	33	68.8	633	2	S62057	proline-rich prote
545	33	68.8	315	2	F98295	hypothetical prote	618	33	68.8	636	2	H96666	probable phytochro
546	33	68.8	321	2	F84597	probable proline-r	619	33	68.8	638	1	QOMSPL	plasma kallikrein
547	33	68.8	325	2	S75095	hypothetical prote	620	33	68.8	647	2	T43952	hypothetical prote
548	33	68.8	343	2	H82171	probable multidrug	621	33	68.8	651	2	T31175	hypothetical prote
549	33	68.8	346	2	H84498	hypothetical prote	622	33	68.8	658	2	T08153	cysteine proteinas
550	33	68.8	347	1	SYECQA	quinolinate synthe	623	33	68.8	659	2	A45184	protein-tyrosine k
551	33	68.8	347	2	AC0593	quinolinate synthe	624	33	68.8	659	2	I49553	protein-tyrosine k
552	33	68.8	347	2	C85577	quinolinate synthe	625	33	68.8	672	2	T32557	hypothetical prote
553	33	68.8	347	2	B99726	inhibin alpha chai	626	33	68.8	682	1	S22700	amphiphysin - chic
554	33	68.8	360	1	A25732	hypothetical prote	627	33	68.8	685	2	T12139	NADH2 dehydrogenas
555	33	68.8	363	2	T16755	hypothetical prote	628	33	68.8	691	2	A25704	synapsin I - rat
556	33	68.8	364	2	F96603	hypothetical prote	629	33	68.8	695	2	T12148	NADH2 dehydrogenas
557	33	68.8	365	2	C37753	quinolinate syntha	630	33	68.8	695	2	S62400	amphiphysin (clone
558	33	68.8	372	2	T01600	hypothetical prote	631	33	68.8	704	2	A30411	synapsin Ia - rat
559	33	68.8	375	2	T48807	hypothetical prote	632	33	68.8	705	2	A35363	synapsin I splice
560	33	68.8	376	2	S71558	probable cell wall	633	33	68.8	706	2	E30411	synapsin Ia - bovi
561	33	68.8	376	2	B96698	unknown protein F1	634	33	68.8	707	2	A46302	PTB-associated spl
562	33	68.8	378	2	T46693	probable 2-nitropr	635	33	68.8	708	1	QQBEB8	UL80 protein - hum
563	33	68.8	379	2	S31719	proline-rich prote	636	33	68.8	714	2	AD2236	p-aminobenzoic aci
564	33	68.8	383	2	AE2295	hypothetical prote	637	33	68.8	724	2	A38749	3-phosphatidylinos
565	33	68.8	384	2	T11580	probable 12-oxophy	638	33	68.8	728	2	T43632	serine/threonine p
566	33	68.8	387	2	T17662	hypothetical prote	639	33	68.8	731	2	T04455	hypothetical prote
567	33	68.8	389	2	B82819	outer membrane pro	640	33	68.8	733	1	S33643	transforming prote
568	33	68.8	390	2	A70656	hypothetical prote	641	33	68.8	736	2	T12217	NADH2 dehydrogenas
569	33	68.8	401	2	T51407	proline-rich prote	642	33	68.8	742	2	F84643	hypothetical prote
570	33	68.8	405	2	T37022	probable lycopene	643	33	68.8	747	1	E71429	probable FCA gamma
571	33	68.8	412	2	E88736	protein F33D4.6a [644	33	68.8	759	2	T44142	DR1 protein [impor
572	33	68.8	417	2	D70547	probable pitA prot	645	33	68.8	765	2	S38099	hypothetical prote
573	33	68.8	419	2	F36788	hypothetical prote	646	33	68.8	767	2	I51083	SOX-L2 - rainbow t
574	33	68.8	419	2	T37577	hypothetical prote	647	33	68.8	768	2	D84271	small heat shock p
575	33	68.8	424	2	T33663	hypothetical prote	648	33	68.8	779	1	S61571	suppressor protein
576	33	68.8	427	2	E72488	probable tryptopha	649	33	68.8	792	2	T43630	serine/threonine p
577	33	68.8	428	2	F81694	pyruvate dehydroge	650	33	68.8	798	2	C96658	probable RNA helic
578	33	68.8	433	2	T02157	hypothetical prote	651	33	68.8	798	2	T19864	hypothetical prote
579	33	68.8	434	2	A75450	conserved hypothet	652	33	68.8	802	2	T21315	hypothetical prote
580	33	68.8	437	2	A88942	protein R13D11.3 [653	33	68.8	807	1	WMAD15	late 100K protein
581	33	68.8	437	2	T32599	hypothetical prote	654	33	68.8	830	2	T10545	hypothetical prote
582	33	68.8	440	2	JC7807	Wiskott-Aldrich sy	655	33	68.8	837	2	T19825	hypothetical prote
583	33	68.8	448	2	A42026	cAMP response elem	656	33	68.8	845	2	E84488	probable RNA helic
584	33	68.8	448	2	T06076	proline-rich prote	657	33	68.8	884	2	T40690	hypothetical prote
585	33	68.8	456	2	B42026	cyclic AMP respons	658	33	68.8	891	2	H75507	hypothetical prote
586	33	68.8	461	2	G86498	CT289 hypothetical	659	33	68.8	892	2	T09071	SH3 domains-contai
587	33	68.8	461	2	G72124	conserved hypothet	660	33	68.8	900	2	A42024	transcription fact
588	33	68.8	464	2	C84686	probable AP2 domai	661	33	68.8	907	2	A24938	hypothetical T2 pr
589	33	68.8	464	2	S22697	extensin - Volvox	662	33	68.8	933	2	S17233	transcription fact
590	33	68.8	474	2	T20041	hypothetical prote	663	33	68.8	957	2	T03829	transcription fact
591	33	68.8	480	2	T10029	hypothetical prote	664	33	68.8	969	2	T02842	probable membrane
592	33	68.8	481	2	JE0377	p70 S6 kinase (EC	665	33	68.8	985	2	T41135	hypothetical prote
593	33	68.8	481	2	T28900	hypothetical prote	666	33	68.8	986	2	T38205	RanBP7/importin-be
594	33	68.8	483	2	A12868	phosphomannomutase	667	33	68.8	998	2	T09492	general transcript
595	33	68.8	493	2	D83005	hypothetical prote	668	33	68.8	1003	2	T34066	hypothetical prote
596	33	68.8	494	2	AC3582	probable blue-copp	669	33	68.8	1005	2	T12546	hypothetical prote
597	33	68.8	496	2	T01564	hypothetical prote	670	33	68.8	1006	2	T42731	atrophin-1 related
598	33	68.8	505	1	S05380	transcription fact	671	33	68.8	1006	2	G86292	hypothetical prote
599	33	68.8	506	2	C97645	phosphomannomutase	672	33	68.8	1015	2	T42013	frequency clock pr
600	33	68.8	511	2	C70803	hypothetical prote	673	33	68.8	1016	2	H71460	probable outer mem
601	33	68.8	512	2	G86914	probable membrane	674	33	68.8	1075	2	T27623	hypothetical prote
602	33	68.8	512	2	B70786	probable integral	675	33	68.8	1080	2	T27622	hypothetical prote
603	33	68.8	524	2	T04564	cytochrome P450 ho	676	33	68.8	1122	2	T14180	exit protein - Myc
604	33	68.8	526	2	T04566	cytochrome P450 ho	677	33	68.8	1151	2	T18535	high molecular mas
605	33	68.8	531	2	B85049	probable transposo	678	33	68.8	1201	2	T29329	hypothetical prote
606	33	68.8	533	2	A85181	FCA delta protein	679	33	68.8	1210	2	I39410	AF-4 protein, spli
607	33	68.8	551	2	S57447	HPBRII-7 protein -	680	33	68.8	1213	2	A58198	serine/proline-ric
608	33	68.8	572	2	G02845	zyxin - human	681	33	68.8	1215	2	T32734	myosin-IA - Acanth
609	33	68.8	574	2	B35149	ipaH protein - Shi	682	33	68.8	1226	2	S69078	hypothetical prote
610	33	68.8	577	2	T28017	hypothetical prote	683	33	68.8	1403	2	S24548	homeotic protein p
611	33	68.8	592	2	T21536	hypothetical prote	684	33	68.8	1536	2	S59841	4-alpha-glucanotra
612	33	68.8	601	2	S56144	SH3 domain binding	685	33	68.8	1556	2	A60988	saliva-interacting
613	33	68.8	620	1	S33253	protein-tyrosine k	686	33	68.8	1561	1	S06839	surface antigen sp

687 33 68.8 1565 2 S04729 surface antigen pa
688 33 68.8 1613 2 JE0272 low density lipopr
689 33 68.8 1613 2 JE0273 low density lipopr
690 33 68.8 1696 2 T00057 hypothetical prote
691 33 68.8 1753 2 S30855 hypothetical prote
692 33 68.8 1872 2 T00339 hypothetical prote
693 33 68.8 1888 2 T14273 zinc finger protei
694 33 68.8 2222 2 A37490 voltage-dependent
695 33 68.8 2251 2 B54972 voltage-dependent
696 33 68.8 2259 2 S29236 calcium channel pr
697 33 68.8 2270 2 A54972 voltage-dependent
698 33 68.8 2272 2 C54972 voltage-dependent
699 33 68.8 2764 2 T13949 neurofibromin - fr
700 33 68.8 2802 2 T13945 neurofibromin - fr
701 33 68.8 2802 2 T13947 neurofibromin - fr
702 33 68.8 3010 1 GNWVTC genome polyprotein
703 33 68.8 3716 2 E70969 probable PPE prote
704 33 68.8 4861 2 S71752 giant protein p619
705 33 68.8 5126 2 S40450 ryanodine receptor
706 32.5 67.7 353 2 G02729 thrombopoietin - h
707 32.5 67.7 353 2 I80105 thrombopoietin pre
708 32.5 67.7 623 2 C84694 probable protein k
709 32.5 67.7 687 2 T08877 Modin - Podospora
710 32.5 67.7 1089 2 T14576 nosA protein - sli
711 32.5 67.7 1456 2 T01397 LTR gag/pol polypr
712 32.5 67.7 2706 2 T28155 variant-specific s
713 32 66.7 51 2 S64676 acetylcholinestera
714 32 66.7 80 2 I65235 testicular luteini
715 32 66.7 101 2 S04804 hypothetical prote
716 32 66.7 101 2 C72540 hypothetical prote
717 32 66.7 108 2 B95314 hypothetical prote
718 32 66.7 118 2 PN0141 lutropin beta chai
719 32 66.7 118 2 PN0139 lutropin beta chai
720 32 66.7 119 2 A61465 lutropin beta chai
721 32 66.7 119 2 D72520 hypothetical prote
722 32 66.7 128 2 T05635 hypothetical prote
723 32 66.7 130 1 ERAD52 early E3B 14.5K pr
724 32 66.7 133 2 JC2202 secretin precursor
725 32 66.7 134 2 D84672 hypothetical prote
726 32 66.7 135 2 H72639 hypothetical prote
727 32 66.7 138 2 S00512 lutropin beta chai
728 32 66.7 139 2 I52320 testicular luteini
729 32 66.7 141 1 UTPGB lutropin beta chai
730 32 66.7 141 1 UTRTB lutropin beta chai
731 32 66.7 141 2 JC4527 luteinizing hormon
732 32 66.7 142 2 S11438 coat protein 2 - b
733 32 66.7 143 1 A60159 interleukin-3 prec
734 32 66.7 145 2 F70919 hypothetical prote
735 32 66.7 147 2 S37485 gene msg1 protein
736 32 66.7 148 2 S39206 proline-rich prote
737 32 66.7 150 2 C72697 hypothetical prote
738 32 66.7 153 2 T14303 heat shock protein
739 32 66.7 154 2 C71162 hypothetical prote
740 32 66.7 157 2 S35037 helix-loop-helix p
741 32 66.7 158 1 ZBBE11 17K zinc-binding p
742 32 66.7 161 2 H85120 probable proline-r
743 32 66.7 166 2 T09593 C1C protein, cold-
744 32 66.7 183 2 I38923 microfibril-associ
745 32 66.7 183 2 A49313 microfibril-associ
746 32 66.7 185 1 JC2394 phospholipase A2 i
747 32 66.7 185 2 D75602 hypothetical prote
748 32 66.7 187 2 E95413 conserved hypothet
749 32 66.7 192 2 F41816 amelogenin (Y chro
750 32 66.7 198 2 T07780 remorin - potato
751 32 66.7 201 2 G70874 probable transcrip
752 32 66.7 204 2 B72733 probable membrane-
753 32 66.7 206 2 T29696 hypothetical prote
754 32 66.7 212 2 T47947 hypothetical prote
755 32 66.7 219 2 T10563 proline-rich prote
756 32 66.7 219 2 E83299 hypothetical prote
757 32 66.7 220 2 S42879 nodulin-30 - kidne
758 32 66.7 226 2 S33024 hypothetical prote
759 32 66.7 229 2 D95398 hypothetical prote

760 32 66.7 233 2 H75359
761 32 66.7 235 1 RWHUT8
762 32 66.7 235 2 F83448 succinate dehydrog
763 32 66.7 235 2 G81138 probable succinate
764 32 66.7 235 2 A83970 hypothetical prote
765 32 66.7 236 2 F82118 succinate dehydrog
766 32 66.7 237 2 A47368 HNF-3/fork head tr
767 32 66.7 238 1 DEEC51 succinate dehydrog
768 32 66.7 238 2 E90722 succinate dehydrog
769 32 66.7 238 2 D85573 succinate dehydrog
770 32 66.7 238 2 AH0136 succinate dehydrog
771 32 66.7 239 2 AC0591 probable proline-r
772 32 66.7 248 2 F84749 nitrate transport
773 32 66.7 248 2 G69329 hypothetical prote
774 32 66.7 248 2 S75632 EspF protein [limpo
775 32 66.7 248 2 F91197 espF [imported] -
776 32 66.7 255 2 B86044 CCAAT-enhancer bin
777 32 66.7 257 2 F81096 conserved hypothet
778 32 66.7 267 1 E64351 nitrate transport
779 32 66.7 268 2 H81844 conserved hypothet
780 32 66.7 268 2 A49303 homeotic protein C
781 32 66.7 270 2 E87649 hypothetical prote
782 32 66.7 271 2 S72782 B1496_F2_81 protei
783 32 66.7 272 2 D83154 probable permease
784 32 66.7 272 2 B75414 molybdenum cofacto
785 32 66.7 276 2 AB3644 hypothetical cytos
786 32 66.7 276 2 B75572 conserved hypothet
787 32 66.7 278 2 B75572 fas ligand - rat
788 32 66.7 278 2 A49266 hypothetical prote
789 32 66.7 278 2 AB2254 hypothetical prote
790 32 66.7 279 2 T10361 probable oligopept
791 32 66.7 283 2 D71130 hypothetical prote
792 32 66.7 284 2 B70834 conserved hypothet
793 32 66.7 284 2 D81075 conserved hypothet
794 32 66.7 284 2 D81866 hypothetical prote
795 32 66.7 284 2 G96815 hypothetical prote
796 32 66.7 286 2 F87220 probable secreted
797 32 66.7 287 2 AD3412 aminomethyltransfe
798 32 66.7 292 2 B83690 aminoglycoside 6-a
799 32 66.7 298 2 D95314 conserved hypothet
800 32 66.7 299 2 AD0058 transcription acti
801 32 66.7 299 2 T17832 hypothetical prote
802 32 66.7 301 2 T33702 hypothetical prote
803 32 66.7 304 2 T16704 hypothetical prote
804 32 66.7 306 2 T06435 ribonuclease S5 ho
805 32 66.7 306 2 T52340 cell wall-plasma m
806 32 66.7 315 2 H83276 probable lipase PA
807 32 66.7 324 2 T22580 hypothetical prote
808 32 66.7 328 2 JQ0985 hydroxyproline-ric
809 32 66.7 334 2 E70519 probable oxidoredu
810 32 66.7 334 2 S20982 chitinase (EC 3.2.
811 32 66.7 338 2 T40774 probable histone p
812 32 66.7 340 2 S32584 structural protein
813 32 66.7 342 2 A96511 unknown protein [i
814 32 66.7 346 2 T51252 dTDPglucose 4,6-de
815 32 66.7 346 2 G87440 rod shape-determin
816 32 66.7 346 2 T27896 hypothetical prote
817 32 66.7 350 2 S22456 hydroxyproline-ric
818 32 66.7 350 2 T14191 extensin homolog T
819 32 66.7 350 2 G75571 MutT/nudix family
820 32 66.7 353 2 E82151 quinolate synthe
821 32 66.7 359 2 I39153 prostanoid DP rece
822 32 66.7 369 2 T22774 hypothetical prote
823 32 66.7 369 2 S20500 hydroxyproline-ric
824 32 66.7 371 2 H86536 CT058 hypothetical
825 32 66.7 371 2 B72087 conserved hypothet
826 32 66.7 374 2 T08747 hypothetical prote
827 32 66.7 384 2 T23604 hypothetical prote
828 32 66.7 391 2 T36311 probable lipase -
829 32 66.7 398 1 WZBEN3 protein-serine/thr
830 32 66.7 399 2 JC6522 paired-type homeod
831 32 66.7 403 2 S52796 prpl2 protein - hu
832 32 66.7 411 2 G86923 probable hydrolase

833 32 66.7 416 1 A31959 lysosome-associate
834 32 66.7 421 2 C82253 folypolyglutamate
835 32 66.7 424 2 S09884 hypothetical prote
836 32 66.7 429 2 H71539 probable dihydroli
837 32 66.7 433 2 T07910 hydroxyproline-ric
838 32 66.7 434 2 F75425 tRNA nucleotidyltr
839 32 66.7 435 2 JN0290 protein-tyrosine k
840 32 66.7 442 2 T16035 hypothetical prote
841 32 66.7 443 2 T17220 hypothetical prote
842 32 66.7 444 2 E83802 hypothetical prote
843 32 66.7 446 2 A34418 H-2 region II bind
844 32 66.7 447 2 T49350 hypothetical prote
845 32 66.7 448 2 D41727 retinoid X recepto
846 32 66.7 449 2 S16748 proline-rich prote
847 32 66.7 451 2 A41651 retinoic acid rece
848 32 66.7 454 1 GQMST1 tumor necrosis fac
849 32 66.7 458 1 KHRZOA oryzain (EC 3.4.22
850 32 66.7 472 2 T24316 hypothetical prote
851 32 66.7 472 2 H70836 hypothetical prote
852 32 66.7 475 2 T32036 hypothetical prote
853 32 66.7 475 2 S47861 mei-1 protein - Ca
854 32 66.7 477 1 I51645 stromelysin 3 (EC
855 32 66.7 477 2 JN0291 protein-tyrosine k
856 32 66.7 482 2 A39285 calcitonin recepto
857 32 66.7 485 2 F87225 conserved membrane
858 32 66.7 489 2 A39257 cytochrome P450 2C
859 32 66.7 490 2 S28166 cytochrome P450 2C
860 32 66.7 490 2 A29782 unspecific monooxy
861 32 66.7 490 2 I48189 cytochrome P450 II
862 32 66.7 490 2 B28516 cytochrome P450 2C
863 32 66.7 496 2 T09300 beta-amylase (EC 3
864 32 66.7 496 2 T11575 beta-amylase (EC 3
865 32 66.7 498 2 I47130 calcitonin recepto
866 32 66.7 503 2 S62018 probable membrane
867 32 66.7 506 2 AG1864 hypothetical prote
868 32 66.7 511 2 A40110 nicotinic acetylch
869 32 66.7 514 2 H89695 protein M03B6.3 [i
870 32 66.7 516 2 H82756 conserved hypothet
871 32 66.7 517 2 E89530 protein H28G03.2 [i
872 32 66.7 520 1 NNEB1T anthranilate synth
873 32 66.7 520 2 AF0653 anthranilate synth
874 32 66.7 520 2 I84718 RXR-beta1 isoform
875 32 66.7 526 2 T33047 hypothetical prote
876 32 66.7 527 2 S13763 protein-tyrosine k
877 32 66.7 533 2 S37781 retinoid X recepto
878 32 66.7 534 2 S21961 proline-rich prote
879 32 66.7 545 2 JC7978 protolin receptor
880 32 66.7 546 2 D83408 hypothetical prote
881 32 66.7 548 2 T41182 probable protein p
882 32 66.7 549 2 A86183 hypothetical prote
883 32 66.7 559 2 I49444 SH3 binding protei
884 32 66.7 561 2 G96752 unknown protein F2
885 32 66.7 570 1 QQBE2R BVRf1 (EC-RF2) pro
886 32 66.7 579 2 T30635 hypothetical prote
887 32 66.7 581 2 I45971 prolactin receptor
888 32 66.7 586 2 A41125 gamma-glutamyltran
889 32 66.7 592 2 D70863 hypothetical prote
890 32 66.7 593 2 S49525 glycoprotein G - s
891 32 66.7 594 2 D64676 oligopeptide ABC t
892 32 66.7 595 2 D71841 hypothetical prote
893 32 66.7 602 2 JU0215 protein-tyrosine k
894 32 66.7 603 2 A75373 probable N-acetylm
895 32 66.7 615 1 KFHU12 coagulation factor
896 32 66.7 620 2 S06733 hydroxyproline-ric
897 32 66.7 624 2 JC5471 regulatory protein
898 32 66.7 628 2 S01955 hypothetical prote
899 32 66.7 630 1 T01380 protein-tyrosine k
900 32 66.7 630 2 T02524 probable RING zinc
901 32 66.7 630 2 I51086 prolactin receptor
902 32 66.7 631 2 I56997 protein-tyrosine k
903 32 66.7 631 2 T31782 hypothetical prote
904 32 66.7 633 2 S47144 mating type A prot
905 32 66.7 634 2 S33339 somatotropin recep

906 32 66.7 636 2 A12552 transposase alr801
907 32 66.7 638 2 A33991 somatotropin recep
908 32 66.7 638 2 S12136 somatotropin recep
909 32 66.7 638 2 B28176 somatotropin recep
910 32 66.7 638 2 A33505 somatotropin recep
911 32 66.7 642 2 T23687 hypothetical prote
912 32 66.7 643 2 T27429 hypothetical prote
913 32 66.7 650 2 AD0352 chaperone protein
914 32 66.7 651 2 JC7705 death receptor-6 -
915 32 66.7 651 2 F86563 hypothetical prote
916 32 66.7 651 2 A39372 potassium channel
917 32 66.7 651 2 A72060 hypothetical prote
918 32 66.7 656 2 S55262 UV-endonuclease -
919 32 66.7 657 2 S32739 Filensin - chicken
920 32 66.7 662 2 I37892 IL12 receptor comp
921 32 66.7 662 2 T17339 hypothetical prote
922 32 66.7 663 2 T40493 hnf-3/forkhead tra
923 32 66.7 668 2 C71868 hypothetical prote
924 32 66.7 676 2 F69394 H+-transporting AT
925 32 66.7 677 2 S33608 somatotropin-bindi
926 32 66.7 680 2 JC5895 killer cell inhibi
927 32 66.7 683 2 T12295 NADH2 dehydrogenas
928 32 66.7 685 2 T18964 hypothetical prote
929 32 66.7 694 2 T12675 NADH2 dehydrogenas
930 32 66.7 695 2 A87494 topoisomerase IV,
931 32 66.7 715 2 B83518 pimeloyl-CoA synth
932 32 66.7 724 2 A38748 3-phosphatidylinos
933 32 66.7 724 2 A38747 phosphatidylinosit
934 32 66.7 731 2 AE1945 hypothetical prote
935 32 66.7 738 2 I49295 IL-12 receptor bet
936 32 66.7 740 2 T12753 NADH2 dehydrogenas
937 32 66.7 747 2 T13683 NADH2 dehydrogenas
938 32 66.7 758 2 T15577 hypothetical prote
939 32 66.7 760 2 T06291 extensin homolog T
940 32 66.7 763 2 JC5234 amine oxidase (cop
941 32 66.7 780 2 T00366 hypothetical prote
942 32 66.7 782 2 T15060 hypothetical prote
943 32 66.7 785 2 S75819 general secretion
944 32 66.7 795 2 T49835 hypothetical prote
945 32 66.7 796 1 A32434 abaA protein - Eme
946 32 66.7 796 2 E96654 hypothetical prote
947 32 66.7 797 2 T33098 hypothetical prote
948 32 66.7 798 2 D96563 probable bZIP prot
949 32 66.7 811 2 S08579 hypothetical prote
950 32 66.7 821 2 T34232 hypothetical prote
951 32 66.7 830 2 I50455 prolactin receptor
952 32 66.7 831 2 JQ1655 prolactin receptor
953 32 66.7 833 2 A90575 conserved hypothet
954 32 66.7 845 2 T52518 related to cytosin
955 32 66.7 845 2 T17291 hypothetical prote
956 32 66.7 848 2 I55498 testicular dynamin
957 32 66.7 873 2 B75514 penicillin-binding
958 32 66.7 886 2 C87031 Cell division prot
959 32 66.7 892 2 T15195 hypothetical prote
960 32 66.7 903 1 VGBEB1 glycoprotein B pre
961 32 66.7 928 2 T38419 Sec76 domain prote
962 32 66.7 945 2 T23965 hypothetical prote
963 32 66.7 948 2 A57640 retinoblastoma bin
964 32 66.7 954 1 S68178 mixed-lineage prot
965 32 66.7 974 1 A40213 optic lobe develop
966 32 66.7 978 2 T03763 BAP-135 protein ho
967 32 66.7 980 2 S54986 regulatory protein
968 32 66.7 990 1 TQEC26 transposase - Esch
969 32 66.7 1003 2 C71139 hypothetical prote
970 32 66.7 1006 2 F87363 medium-chain-fatty
971 32 66.7 1023 2 E71376 conserved hypothet
972 32 66.7 1027 2 F82829 acriflavin resista
973 32 66.7 1039 2 T22117 hypothetical prote
974 32 66.7 1046 2 F71432 hypothetical prote
975 32 66.7 1070 2 T25591 hypothetical prote
976 32 66.7 1087 2 S58147 protein kinase - f
977 32 66.7 1097 1 S47220 protein kinase C (

978 32 66.7 1102 2 T28666 protein kinase C-r

979 32 66.7 1104 2 I38869 transcription fact
980 32 66.7 1106 1 A39299 DNA-directed DNA p
981 32 66.7 1107 1 JQ1658 genome polyprotein
982 32 66.7 1119 2 T16720 hypothetical prote
983 32 66.7 1122 2 T47424 hypothetical prote
984 32 66.7 1128 2 AH1949 hypothetical prote
985 32 66.7 1136 1 S57845 protein-tyrosine k
986 32 66.7 1137 2 A86335 T20H2.9 protein -
987 32 66.7 1141 2 T29185 hypothetical prote
988 32 66.7 1160 2 S34809 avrBs3-2 protein -
989 32 66.7 1173 2 T25985 hypothetical prote
990 32 66.7 1187 2 JQ0316 hypothetical 125K
991 32 66.7 1199 2 T15826 hypothetical prote
992 32 66.7 1201 2 G86441 unknown protein [i
993 32 66.7 1207 2 T00378 KIAA0641 protein -
994 32 66.7 1212 2 S27771 RNA-directed DNA p
995 32 66.7 1240 2 T48800 SMT4 related prote
996 32 66.7 1271 2 S37958 myosin heavy chain
997 32 66.7 1281 2 JC5368 dynactin 1 - mouse
998 32 66.7 1290 2 T00018 period protein hom
999 32 66.7 1291 2 T00019 period protein hom
1000 32 66.7 1308 2 T15280 hypothetical prote

ALIGNMENTS

RESULT 1
A56031
potassium channel KCNAs - human
N;/Alternate names: potassium channel HK2; potassium channel PCN1; shaker-related potassi
C;/Species: Homo sapiens (man)
C;/Date: 05-Apr-1995 #sequence_revision 05-Apr-1995 #text_change 09-Jul-2004
C;/Accession: A56031; A38556; B39922; A38074
R;/Phillipson, L.H.; LaMendola, J.; Bell, G.I.; Steiner, D.F.
submitted to GenBank, September 1990
A;/Reference number: A56031
A;/Accession: A56031
A;/Molecule type: mRNA
A;/Residues: 1-613 <PHI>
A;/Cross-references: UNIPROT:P22460; UNIPARC:UPI000016AE76; GB:M55513; NID:G189653; PIDN:
R;/Phillipson, L.H.; Hice, R.E.; Schaefer, K.; LaMendola, J.; Bell, G.I.; Nelson, D.J.; St
Proc. Natl. Acad. Sci. U.S.A. 88, 53-57, 1991
A;/Title: Sequence and functional expression in Xenopus oocytes of a human insulinoma and
A;/Reference number: A38556; MUID:91095456; PMID:1986382
A;/Accession: A38556
A;/Status: nucleic acid sequence not shown
A;/Molecule type: DNA
A;/Residues: 1-56,'G',58-137,'L',139-213,'R',215-227,'P',229-545,'QG',546-613 <PH2>
A;/Cross-references: UNIPARC:UPI00001779F3; GB:M55513
R;/Tamkun, M.M.; Knoth, K.M.; Walbridge, J.A.; Kroemer, H.; Roden, D.M.; Glover, D.M.
FASEB J. 5, 331-337, 1991
A;/Title: Molecular cloning and characterization of two voltage-gated K(+) channel cDNAs
A;/Reference number: A39922; MUID:91160866; PMID:2001794
A;/Accession: B39922
A;/Status: preliminary
A;/Molecule type: mRNA
A;/Residues: 1-54,56-137,'L',139-186,'G',189-213,'R',215-227,'P',229-297,'PTQRGH',309-558
A;/Cross-references: UNIPARC:UPI00001779F4; GB:M60451
R;/Curran, M.E.; Landes, G.M.; Keating, M.T.
Genomics 12, 729-737, 1992
A;/Title: Molecular cloning, characterization, and genomic localization of a human potase
A;/Reference number: A38074; MUID:92241872; PMID:1349297
A;/Accession: A38074
A;/Molecule type: DNA
A;/Residues: 1-137,'L',139-153,'R',155-213,'R',215-227,'P',229-281,'V',283-578,'QLPPREV'
A;/Cross-references: UNIPARC:UPI000016AEEA; GB:M83254; NID:G190202; PIDN:AAA60146.1; PID:
A;/Experimental source: heart
A;/Note: sequence extracted from NCBI backbone (NCBIN:98573, NCBIP:98577)
C;/Genetics:
A;/Gene: GDB:KCNAS
A;/Cross-references: GDB:127904; OMIM:176267
A;/Map position: 12p13.33-12p13.31

C;/Superfamily: potassium channel protein drk1
C;/Keywords: glycoprotein; phosphoprotein; potassium channel; transmembrane protein; volt
F;125,190/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;557/Binding site: phosphate (Ser) (covalent) #status predicted
Query Match 97.9%; Score 47; DB 2; Length 613;
Best Local Similarity 88.9%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 RLPPLPXP 12
Db ||||| |
65 RLPPLPDP 73
RESULT 2
T07173
hypothetical protein SEND32, senescence down-regulated - tomato (fragment)
C;/Species: Lycopersicon esculentum (tomato)
C;/Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
C;/Accession: T07173
R;/John, I.; Hackett, R.; Cooper, W.; Drake, R.; Farrell, A.; Grierson, D.
Plant Mol. Biol. 33, 641-651, 1997
A;/Title: Cloning and characterization of tomato leaf senescence-related cDNAs.
A;/Reference number: Z15976; MUID:97238471; PMID:9132056
A;/Accession: T07173
A;/Status: preliminary; translated from GB/EMBL/DBJ
A;/Molecule type: mRNA
A;/Residues: 1-162 <JOH>
A;/Cross-references: UNIPROT:Q43516; UNIPARC:UPI000009DEBF; EMBL:Z75519; NID:G1418979; PI
A;/Experimental source: cultivar Ailsa Craig; leaf; clone SEND32
C;/Superfamily: herpesvirus immediate-early protein IE110; RING finger homology

Query Match 89.6%; Score 43; DB 2; Length 162;
Best Local Similarity 87.5%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 5 PLPPLPXP 12
Db ||||| |
155 PLPPLPSP 162
RESULT 3
E89008
protein W08A12.3 [imported] - Caenorhabditis elegans
C;/Species: Caenorhabditis elegans
C;/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;/Accession: E89008
R;/anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;/Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A;/Reference number: A75000; MUID:99069613; PMID:9851916
A;/Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/projects/C_ele
A;/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;/Accession: E89008
A;/Status: preliminary
A;/Molecule type: DNA
A;/Residues: 1-198 <STO>
A;/Cross-references: UNIPROT:O01621; UNIPARC:UPI0000076A91; GB:chr_V; PIDN:AABS2319.1; PI
C;/Genetics:
A;/Gene: W08A12.3
A;/Map position: 5

Query Match 87.5%; Score 42; DB 2; Length 198;
Best Local Similarity 87.5%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 5 PLPPLPXP 12
Db ||||| |
98 PLPPLPEP 105

RESULT 4
I38707

Fas ligand - human
C;Species: Homo sapiens (man)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 09-Jul-2004
C;Accession: I38707; JC2340; S57565; I38554
R;Takahashi, T.; Tanaka, M.; Inazawa, J.; Abe, T.; Suda, T.; Nagata, S.
Int. Immunol. 6, 1567-1574, 1994
A;Title: Human Fas ligand: gene structure, chromosomal location and species specificity.
A;Reference number: I38707; MUID:95127560; PMID:7826947
A;Accession: I38707
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-281 <RES>
A;Cross-references: UNIPROT:P48023; UNIPARC:UPI000000D91A; EMBL:U11821; NID:G595430; PID
R;Mita, E.; Hayashi, N.; Iio, S.; Takehara, T.; Hijioka, T.; Kasahara, A.; Fusamoto, H.;
Biochem. Biophys. Res. Commun. 204, 468-474, 1994
A;Title: Role of Fas ligand in apoptosis induced by hepatitis C virus infection.
A;Reference number: JC2340; MUID:95071350; PMID:7980502
A;Accession: JC2340
A;Molecule type: DNA
A;Residues: 1-281 <MIT>
A;Cross-references: UNIPARC:UPI000000D91A; GB:D38122; DDBJ:D29820; NID:G601892; PIDN:BAA
R;Schatzlein, C.E.
submitted to the EMBL Data Library, June 1995
A;Reference number: S57565
A;Accession: S57565
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-281 <SCH>
A;Cross-references: UNIPARC:UPI000000D91A; EMBL:X89102; NID:G887455; PID:G887456
R;Alderson, M.R.; Tough, T.W.; Davis-Smith, T.; Braddy, S.; Falk, B.; Schooley, K.A.; G
J. Exp. Med. 181, 71-77, 1995
A;Title: Fas ligand mediates activation-induced cell death in human T lymphocytes.
A;Reference number: I38554; MUID:95105731; PMID:7528780
A;Accession: I38554
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-281 <RE2>
A;Cross-references: UNIPARC:UPI000000D91A; EMBL:U08137; NID:G624627; PIDN:AAC50071.1; PI
C;Genetics:
A;Gene: FasL
A;Introns: 151/1; 116/3
C;Keywords: glycoprotein; transmembrane protein
F;80-102/Domain: transmembrane #status predicted <TMM>
F;76,184,250,260/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 87.5%; Score 42; DB 2; Length 281;
Best Local Similarity 87.5%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLPXP 12
|||
Db 62 PLPPLPLP 69

RESULT 5
S50755
hypothetical protein VSP-3 - Chlamydomonas reinhardtii
C;Species: Chlamydomonas reinhardtii
C;Date: 14-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
C;Accession: S50755
R;Woessner, J.P.; Molendijk, A.J.; van Egmond, P.; Klis, F.M.; Goodenough, U.W.; Haring,
Plant Mol. Biol. 26, 947-960, 1994
A;Title: Domain conservation in several volvocalean cell wall proteins.
A;Reference number: S50754; MUID:95093034; PMID:8000007
A;Accession: S50755
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-473 <WOE>
A;Cross-references: UNIPROT:Q39620; UNIPARC:UPI000009DF5C; EMBL:L29029; NID:G530875; PID

Query Match 87.5%; Score 42; DB 2; Length 473;
Best Local Similarity 77.8%; Pred. No. 69;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
|||
Db 360 RPSPLPSP 368

RESULT 6
JC7509
glycoprotein VI-1 - human
C;Species: Homo sapiens (man)
C;Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 09-Jul-2004
C;Accession: JC7509; PC7101
R;Ezumi, Y.; Uchiyama, T.; Takayama, H.
Biochem. Biophys. Res. Commun. 277, 27-36, 2000
A;Title: Molecular cloning, genomic structure, chromosomal localization, and alternative
A;Reference number: JC7509; MUID:20483673; PMID:11027634
A;Contents: Platelet
A;Accession: JC7509
A;Molecule type: mRNA
A;Residues: 1-339 <EZU>
A;Cross-references: UNIPROT:Q9UIF2; UNIPARC:UPI000006F4A8; DDBJ:AB043819
A;Accession: PC7101
A;Molecule type: protein
A;Residues: 28-41;62-79;114-142 <EZ2>
A;Cross-references: UNIPARC:UPI000017A509; UNIPARC:UPI000017A50A; UNIPARC:UPI000017A50B
C;Comment: This protein, which belongs to the immunoglobulin superfamily, is the major c
or gamma chain as a signal transducing subunit, and plays some roles in cancer cells.
C;Genetics:
A;Gene: gpVI-1
A;Map position: 19q13.4
A;Introns: 62/1; 95/1; 353/1; 638/1; 692/1; 752/1; 803/1
C;Keywords: glycoprotein; immunoglobulin; platelet

Query Match 85.4%; Score 41; DB 2; Length 339;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RPLPPLP 10
|||
Db 307 RPLPPLP 313

RESULT 7
D88088
protein B0454.1 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C;Accession: D88088
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_eleg
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: D88088
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-348 <STO>
A;Cross-references: UNIPARC:UPI000017A57E; GB:chr_II; PIDN:AB70947.1; PID:G2429396; GSPD
C;Genetics:
A;Gene: B0454.1
A;Map position: 2

Query Match 85.4%; Score 41; DB 2; Length 348;
Best Local Similarity 87.5%; Pred. No. 69;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLPXP 12
|||
Db 37 PLPPLPPP 44

RESULT 8

S03116
Gene 33 protein, hepatic - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 28-Feb-1990 #sequence revision 28-Feb-1990 #text_change 09-Jul-2004
C;Accession: S03116; S03402; B30568; S08283
R;Chrapkiewicz, N.B.; Davis, C.M.; Chu, D.T.W.; Caldwell, C.M.; Granner, D.K.
Nucleic Acids Res. 17, 6651-6667, 1989
A;Title: Rat gene 33: analysis of its structure, messenger RNA and basal promoter activation
A;Reference number: S03116; MUID:89385990; PMID:2780291
A;Accession: S03116
A;Molecule type: DNA
A;Residues: 1-459 <CHR>
A;Cross-references: UNIPROT:P05432; UNIPARC:UPI000012F0FD; EMBL:X07266; NID:g57568; PIDN:R;Lee, K.L.; Makkinje, A.; Ch'ang, L.Y.; Kenney, F.T.
Arch. Biochem. Biophys. 269, 106-113, 1989
A;Title: Molecular cloning and analysis of full-length cDNAs cognate to a rat gene under
A;Reference number: S03402; MUID:89133523; PMID:2916834
A;Accession: S03402
A;Molecule type: mRNA
A;Residues: 1-17, 'K', 19-66, 143-191, 'T', 193-301, 'L', 303-310, 'L', 312-395, 'L', 397-409, 'L', 4
A;Cross-references: UNIPARC:UPI000017C912
R;Lee, K.L.; Makkinje, A.; Ch'ang, L.Y.; Kenney, F.T.
Arch. Biochem. Biophys. 276, 554, 1990
A;Reference number: S08283
A;Contents: annotation
A;Note: this is a revision of the nucleotide translation of residues 18, 192, 302, 311,
C;Genetics:
A;Gene: 33
C;Keywords: alternative splicing; liver
F;1-459/Product: gene 33 protein, long form #status predicted <MAT1>
F;1-66,143-459/Product: gene 33 protein, short form #status predicted <MAT2>

Query Match 85.4%; Score 41; DB 2; Length 459;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RLPPLP 10
Db 150 RLPPLP 156

RESULT 9
T46917
hypothetical protein DKFZp762K137.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 09-Jul-2004
C;Accession: T46917
R;Ottenwaelder, B.; Obermaier, B.; Mewes, H.W.; Weil, B.; Wiemann, S.
submitted to the Protein Sequence Database, February 2000
A;Reference number: Z24136
A;Accession: T46917
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-477 <AAA>
A;Cross-references: UNIPROT:Q9NSQ8; UNIPARC:UPI000006CEC3; EMBL:AL157482
A;Experimental source: adult melanoma (MeWo cell line); clone DKFZp762K137
C;Genetics:
A;Note: DKFZp762K137.1

Query Match 85.4%; Score 41; DB 2; Length 477;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RLPPLP 10
Db 228 RLPPLP 234

RESULT 10

A46101
protein-tyrosine-phosphatase (EC 3.1.3.48) nonreceptor type PTP61F, short splice form -
C;Species: Drosophila melanogaster
C;Date: 08-May-1995 #sequence_revision 12-May-1995 #text_change 09-Jul-2004
C;Accession: A46101
R;McLaughlin, S.; Dixon, J.E.
J. Biol. Chem. 268, 6839-6842, 1993
A;Title: Alternative splicing gives rise to a nuclear protein tyrosine phosphatase in Drosophila
A;Reference number: A46101; MUID:93216607; PMID:8463208
A;Accession: A46101
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-535 <MCL>
A;Cross-references: UNIPROT:Q9W0G1; UNIPARC:UPI000016BD13; GB:L11251; NID:g290265; PIDN:R;McLaughlin, S.; Dixon, J.E.
A;Note: authors translated the codon TTC for residue 382 as Ile, and CGA for residue 528
C;Genetics:
A;Gene: FlyBase:Ptp61F
A;Cross-references: FlyBase:FBgn0003138
C;Superfamily: protein-tyrosine-phosphatase, nonreceptor type PTP61F; protein-tyrosine-phosphatase, nonreceptor type PTP61F; phosphoric monoester hydrolase; tyrosine phosphatase
C;Keywords: alternative splicing; phosphoprotein; phosphoric monoester hydrolase; tyrosine phosphatase
F;60-285/Domain: protein-tyrosine-phosphatase homology <PTP>
F;237/Active site: Cys (phosphocysteine intermediate) #status predicted
F;243/Binding site: substrate phosphate (Arg) #status predicted

Query Match 85.4%; Score 41; DB 2; Length 535;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RLPPLP 10
Db 391 RLPPLP 397

RESULT 11

B46101
protein-tyrosine-phosphatase (EC 3.1.3.48), nonreceptor type PTP61F, long splice form -
C;Species: Drosophila melanogaster
C;Date: 08-May-1995 #sequence_revision 12-May-1995 #text_change 09-Jul-2004
C;Accession: B46101
R;McLaughlin, S.; Dixon, J.E.
J. Biol. Chem. 268, 6839-6842, 1993
A;Title: Alternative splicing gives rise to a nuclear protein tyrosine phosphatase in Drosophila
A;Reference number: A46101; MUID:93216607; PMID:8463208
A;Accession: B46101
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-548 <MCL>
A;Cross-references: UNIPROT:Q9W0G1; UNIPARC:UPI000016BD12; GB:L11251
A;Note: authors translated the codon TTC for residue 382 as Ile
C;Genetics:
A;Gene: FlyBase:Ptp61F
A;Cross-references: FlyBase:FBgn0003138
C;Superfamily: protein-tyrosine-phosphatase, nonreceptor type PTP61F; protein-tyrosine-phosphatase, nonreceptor type PTP61F; phosphoric monoester hydrolase; tyrosine phosphatase
C;Keywords: alternative splicing; phosphoprotein; phosphoric monoester hydrolase; tyrosine phosphatase
F;60-285/Domain: protein-tyrosine-phosphatase homology <PTP>
F;237/Active site: Cys (phosphocysteine intermediate) #status predicted
F;243/Binding site: substrate phosphate (Arg) #status predicted

Query Match 85.4%; Score 41; DB 2; Length 548;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RLPPLP 10
Db 391 RLPPLP 397

RESULT 12

T26064
hypothetical protein W01G7.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T26064
R;Percy, C.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z20146
A;Accession: T26064
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-627 <WIL>
A;Cross-references: UNIPROT:Q9XVH7; UNIPARC:UPI0000079E8B; EMBL:Z81135; PIDN:CAB03453.1;
A;Experimental source: clone W01G7
C;Genetics:
A;Gene: CESP:W01G7.1
A;Map position: 2
A;Introns: 82/2; 121/1; 182/2; 588/3
C;Superfamily: Caenorhabditis elegans hypothetical protein W01G7.1

Query Match 85.4%; Score 41; DB 2; Length 627;
Best Local Similarity 87.5%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 PLPPLPXP 12
|||||||
Db 513 PLPPLPPP 520

RESULT 13
TVMVRR
protein-tyrosine kinase (EC 2.7.1.112) fgr - feline sarcoma virus (strain Gardner-Rasheed)
C;Species: feline sarcoma virus
A;Note: host Felis sp. (cat)
C;Date: 27-Nov-1985 #sequence_revision 26-May-1995 #text_change 09-Jul-2004
C;Accession: A00653; A03937
R;Naharro, G.; Robbins, K.C.; Reddy, E.P.
Science 223, 63-66, 1984
A;Title: Gene product of v-fgr onc: hybrid protein containing a portion of actin and a b
A;Reference number: A00653; MUID:84097512; PMID:6318314
A;Accession: A00653
A;Molecule type: DNA
A;Residues: 1-663 <NAH>
A;Cross-references: UNIPROT:P00544; UNIPARC:UPI000017101E; GB:X00255; GB:K01487; NID:g61
A;Note: the authors translated the codon GAT for residue 14 as Glu
C;Comment: This protein is synthesized as a gag-fgr polypeptide.
C;Genetics:
A;Gene: fgr
C;Superfamily: feline sarcoma virus protein-tyrosine kinase fgr; protein kinase homology
C;Keywords: ATP; autophosphorylation; oncogene; phosphoprotein; phosphotransferase; poly
F;1-118/Region: gag polypeptide similarity
F;141-268/Region: actin similarity
F;285-382/Domain: SH2 homology <SH2>
F;402-660/Domain: protein kinase homology <KIN>
F;410-418/Region: protein kinase ATP-binding motif
F;432/Active site: Lys #status predicted

Query Match 85.4%; Score 41; DB 1; Length 663;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RPLPPLP 10
|||||||
Db 131 RPLPPLP 137

RESULT 14
F84870
hypothetical protein At2g43800 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: F84870
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: F84870
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-894 <STO>
A;Cross-references: UNIPROT:O22824; UNIPARC:UPI00000A2AE1; GB:AE002093; NID:g2281090; PI:
C;Genetics:
A;Gene: At2g43800
A;Map position: 2

Query Match 85.4%; Score 41; DB 2; Length 894;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RPLPPLP 10
|||||||
Db 289 RPLPPLP 295

RESULT 15
T16346
hypothetical protein F42C5.10 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T16346
R;Du, Z.
submitted to the EMBL Data Library, November 1995
A;Description: The sequence of C. elegans cosmid F42C5.
A;Reference number: Z18497
A;Accession: T16346
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1225 <DUZ>
A;Cross-references: UNIPROT:Q20330; UNIPARC:UPI000017B9B3; EMBL:U40799; NID:g1065935; PI:
C;Genetics:
A;Gene: CESP:F42C5.10
A;Introns: 11/3; 280/2; 312/3; 646/1; 786/1; 869/1; 1034/1; 1080/2; 1108/1; 1170/1

Query Match 85.4%; Score 41; DB 2; Length 1225;
Best Local Similarity 87.5%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 PLPPLPXP 12
|||||||
Db 491 PLPPLPPP 498

RESULT 16
T36365
proline-rich protein SCE94.05 - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T36365
R;Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, April 1999
A;Reference number: Z21573
A;Accession: T36365
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-134 <OLI>
A;Cross-references: UNIPROT:Q9X8L8; UNIPARC:UPI00000DAFBA; EMBL:AL049628; PIDN:CAB40854.
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SCE94.05

Query Match 83.3%; Score 40; DB 2; Length 134;
Best Local Similarity 77.8%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RPLPPLPXP 12
|||||||
Db 102 RLPSPPPP 110

```
RESULT 17
GNVQL2
genome-linked protein - potato leaf roll virus
C;Species: potato leaf roll virus
A;Note: host Solanum tuberosum (potato)
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
C;Accession: JQ0002; B60080
R;Kawchuk, L.M.; Martin, R.R.; Rochon, D.M.; McPherson, J.
J. Gen. Virol. 70, 783-788, 1989
A;Title: Identification and characterization of the potato leafroll virus putative coat
A;Reference number: A92803; MUID:89279259; PMID:2732704
A;Accession: JQ0002
A;Molecule type: mRNA
A;Residues: 1-156 <KAW>
A;Cross-references: UNIPROT:P17523; UNIPARC:UPI00000005FE; GB:D13753; NID:g222290; PIDN:
R;Tacke, E.; Sarkar, S.; Salamini, F.; Rohde, W.
Arch. Virol. 105, 153-163, 1989
A;Title: Cloning of the gene for the capsid protein of potato leafroll virus.
A;Reference number: A60080; MUID:89321807; PMID:2751428
A;Accession: B60080
A;Molecule type: mRNA
A;Residues: 1-9,'G',11-40,'V',42-120,'A',122-156 <TAC>
A;Cross-references: UNIPARC:UPI000002D3ED
C;Superfamily: potato leaf roll virus genome-linked protein
C;Keywords: genome-linked protein

Query Match      83.3%; Score 40; DB 1; Length 156;
Best Local Similarity 77.8%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      4 RPLPPLPXP 12
      |||||
Db      114 RPLPPPPVP 122

RESULT 18
GNVQLL
genome-linked protein - potato leaf roll virus (strain 1)
C;Species: potato leaf roll virus
A;Note: host Solanum tuberosum (potato)
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
C;Accession: JA0121; S36689
R;Mayo, M.A.; Robinson, D.J.; Jolly, C.A.; Hyman, L.
J. Gen. Virol. 70, 1037-1051, 1989
A;Title: Nucleotide sequence of potato leafroll luteovirus RNA.
A;Reference number: JA0119; MUID:89279282; PMID:2732710
A;Accession: JA0121
A;Molecule type: genomic RNA
A;Residues: 1-156 <MAY>
A;Cross-references: UNIPROT:P17524; UNIPARC:UPI0000138CCE; EMBL:X14600; NID:g222293; PID
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1989
C;Superfamily: potato leaf roll virus genome-linked protein
C;Keywords: genome-linked protein

Query Match      83.3%; Score 40; DB 1; Length 156;
Best Local Similarity 77.8%; Pred. NO. 41;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      4 RPLPPLPXP 12
      |||||
Db      114 RPLPPPPVP 122

RESULT 19
GNVQWA
genome-linked protein - potato leaf roll virus (strain Wageningen)
N;Alternate names: VPg protein
C;Species: potato leaf roll virus
A;Note: host Solanum tuberosum (potato)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C;Accession: S03550
```

```
R;van der Wilk, F.; Huisman, M.J.; Cornelissen, B.J.C.; Huttinga, H.; Goldbach, R.
FEBS Lett. 245, 51-56, 1989
A;Title: Nucleotide sequence and organization of potato leafroll virus genomic RNA.
A;Reference number: S03546; MUID:89171329; PMID:2466700
A;Accession: S03550
A;Molecule type: genomic RNA
A;Residues: 1-156 <VAN>
A;Cross-references: UNIPROT:P11625; UNIPARC:UPI0000138CCF; EMBL:Y07496; NID:g61198; PIDN
C;Superfamily: potato leaf roll virus genome-linked protein
C;Keywords: genome-linked protein

Query Match      83.3%; Score 40; DB 1; Length 156;
Best Local Similarity 77.8%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      4 RPLPPLPXP 12
      |||||
Db      114 RPLPPPPVP 122

RESULT 20
T15811
hypothetical protein C46H3.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T15811
R;Favello, T.
submitted to the EMBL Data Library, November 1995
A;Description: The sequence of C. elegans cosmid C46H3.
A;Reference number: Z18409
A;Accession: T15811
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-215 <FAV>
A;Cross-references: UNIPROT:Q18661; UNIPARC:UPI000007AB2C; EMBL:U41271; NID:gl086855; PII
C;Genetics:
A;Gene: CESP:C46H3.1
A;Introns: 19/1; 71/1; 99/1; 126/1

Query Match      83.3%; Score 40; DB 2; Length 215;
Best Local Similarity 75.0%; Pred. No. 57;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      5 PLPPLPXP 12
      |||||
Db      104 PLPPMPVP 111

RESULT 21
T16669
hypothetical protein R03H10.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 04-Mar-2000
C;Accession: T16669
R;Wilcox, L.
submitted to the EMBL Data Library, June 1995
A;Description: The sequence of C. elegans cosmid R03H10.
A;Reference number: Z18557
A;Accession: T16669
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-224 <WIL>
A;Cross-references: UNIPARC:UPI000017941B; EMBL:U29382; NID:g868226; PID:g868235; PIDN:A
A;Experimental source: strain Bristol N2
C;Genetics:
A;Gene: CESP:R03H10.2
A;Introns: 47/2; 68/3; 127/2; 159/2; 179/2; 209/2
C;Superfamily: Caenorhabditis elegans hypothetical protein R03H10.2

Query Match      83.3%; Score 40; DB 2; Length 224;
Best Local Similarity 75.0%; Pred. No. 60;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

Qy 5 PLPPLXP 12
Db 164 PLPPVPSP 171

RESULT 22
T29543
hypothetical protein F48C1.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T29543
R:Gattung, S.; Le, T.T.
submitted to the EMBL Data Library, April 1997
A:Description: The sequence of C. elegans cosmid F48C1.
A:Reference number: 220638
A:Accession: T29543
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-245 <GAT>
A:Cross-references: UNIPROT:O01572; UNIPARC:UPI000007781C; EMBL:U97015; PIDN:AAB52342.1;
A:Experimental source: strain Bristol N2; clone F48C1
C:Genetics:
A:Gene: CESP:F48C1.6
A:Map position: 1
A:Introns: 52/3; 96/2; 123/3; 161/3
C:Superfamily: Caenorhabditis elegans hypothetical protein F48C1.6

Query Match 83.3%; Score 40; DB 2; Length 245;
Best Local Similarity 66.7%; Pred. No. 66;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RPLPLXP 12
Db 208 RSPPIPPSP 216

RESULT 23
C70551
hypothetical protein Rv2507 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: C70551
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: C70551
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-273 <COL>
A:Cross-references: UNIPROT:O06170; UNIPARC:UPI00001652FA; GB:Z955556; GB:AL123456; NID:9
A:Experimental source: strain H37Rv
C:Genetics:
A:Gene: Rv2507

Query Match 83.3%; Score 40; DB 2; Length 273;
Best Local Similarity 75.0%; Pred. No. 73;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 PLPPLXP 12
Db 144 PLPPMPGP 151

RESULT 24
T41395
probable dna polymerase alpha-primase associated subunit - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Dec-2004
C:Accession: T41395

R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Murphy, L.; Harris, D.
submitted to the EMBL Data Library, May 1998
A:Reference number: Z21991
A:Accession: T41395
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-574 <WOO>
A:Cross-references: UNIPROT:O74946; UNIPARC:UPI000012977E; EMBL:AL023704; PIDN:CAA19261.
A:Experimental source: strain 972h-; cosmid c553
C:Genetics:
A:Gene: SPDB:SPCC553.09C
A:Map position: 3
A:Introns: 89/2; 415/3; 518/3
C:Superfamily: DNA polymerase alpha, subunit B

Query Match 83.3%; Score 40; DB 2; Length 574;
Best Local Similarity 75.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 PLPPLXP 12
Db 278 PIPPLPP 285

RESULT 25
C81594
hypothetical protein CP0282 [imported] - Chlamydophila pneumoniae (strain AR39)
C:Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: C81594
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, J.; C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: A81500; MUID:20150255; PMID:10684935
A:Accession: C81594
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-775 <REA>
A:Cross-references: UNIPROT:Q9JS20; UNIPARC:UPI00000D2F9E; GB:AE002161; NID
A:Experimental source: strain AR39, HL cells
C:Genetics:
A:Gene: CP0282

Query Match 83.3%; Score 40; DB 2; Length 775;
Best Local Similarity 75.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 PLPPLXP 12
Db 599 PLPPVPSP 606

RESULT 26
D86549
hypothetical protein CPj0472 [imported] - Chlamydophila pneumoniae (strain J138)
C:Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: D86549
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A:Reference number: A86491; MUID:20330349; PMID:10871362
A:Accession: D86549
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-775 <STO>
A:Cross-references: UNIPROT:Q9JS20; UNIPARC:UPI00000D2F9E; GB:BA000008; NID:g8978842; PI
A:Experimental source: strain J138
C:Genetics:
A:Gene: CPj0472

Query Match 83.3%; Score 40; DB 2; Length 775;

Best Local Similarity 75.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLPXP 12
|||:|
Db 599 PLPPVPSP 606

RESULT 27
B72074
hypothetical protein - Chlamydophila pneumoniae (strain CWL029)
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: B72074
R;Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A;Title: Comparative genomes of Clamydia pneumoniae and C. trachomatis.
A;Reference number: A72000; MUID:99206606; PMID:10192388
A;Accession: B72074
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-775 <ARN>
A;Cross-references: UNIPROT:Q9Z879; UNIPARC:UPI00000470F0; GB:AE001632; GB:AE001363; NID
A;Experimental source: strain CWL029
C;Genetics:
A;Gene: CPn0472

Query Match 83.3%; Score 40; DB 2; Length 775;
Best Local Similarity 75.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLPXP 12
|||:|
Db 599 PLPPVPSP 606

RESULT 28
S60735
splicing factor SF3a 120K chain - human
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S60735; S60733
R;Kraemer, A.; Mulhauser, F.; Wersig, C.; Groening, K.; Bilbe, G.
RNA 1, 260-272, 1995
A;Title: Mammalian splicing factor SF3a120 represents a new member of the SURP family of
A;Reference number: S60733; MUID:96079958; PMID:7489498
A;Accession: S60735
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-793 <KRA>
A;Cross-references: UNIPROT:Q15459; UNIPARC:UPI0000000C88; EMBL:X85237; NID:G899297; PID
A;Accession: S60733
A;Molecule type: protein
A;Residues: 51-62;82-94;270-275;397-414;448-463 <KRA2>
A;Cross-references: UNIPARC:UPI0000174C3E; UNIPARC:UPI0000174C3F; UNIPARC:UPI0000174C40;
C;Genetics:
A;Gene: GDB:SF3A120; PRP21; SAP114
A;Cross-references: GDB:9955873
A;Map position: 22q12.1-22qter
C;Superfamily: human splicing factor SF3a 120K chain; ubiquitin homology
C;Keywords: pre-mRNA splicing
F;714-790/Domain: ubiquitin homology <UBH>

Query Match 83.3%; Score 40; DB 1; Length 793;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLPXP 12
:|||||
Db 399 KPLPPAPAP 407

.RESULT 29

G30010
hypothetical ORF-8 protein - Leishmania tarentolae mitochondrion
C;Species: mitochondrion Leishmania tarentolae
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 07-Dec-1999
C;Accession: G30010
R;de la Cruz, V.F.; Neckelmann, N.; Simpson, L.
J. Biol. Chem. 259, 15136-15147, 1984
A;Title: Sequences of six genes and several open reading frames in the kinetoplast maxic.
A;Reference number: A2848; MUID:85079995; PMID:6096360
A;Accession: G30010
A;Molecule type: DNA
A;Residues: 1-93
A;Cross-references: UNIPARC:UPI000017B58D; GB:M10126
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC6
C;Keywords: mitochondrion

Query Match 81.2%; Score 39; DB 2; Length 93;
Best Local Similarity 75.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 PLPPLPXP 12
|||
Db 43 PLPPFPSP 50

RESULT 30
F83532
conserved hypothetical protein PA0915 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: F83532
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br.
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A;Reference number: A82950; MUID:2043737; PMID:10984043
A;Accession: F83532
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-153 <STO>
A;Cross-references: UNIPROT:Q9I542; UNIPARC:UPI000000C51B6; GB:AE004525; GB:AE004091; NID
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA0915

Query Match 81.2%; Score 39; DB 2; Length 153;
Best Local Similarity 77.8%; Pred. No. 55;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLPXP 12
|||||
Db 75 RPLPPLP 83

RESULT 31
T19341
hypothetical protein Cl6D6.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T19341
R;Gardner, A.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19111
A;Accession: T19341
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-173 <WIL>
A;Cross-references: UNIPROT:O62063; UNIPARC:UPI000007A889; EMBL:Z81472; PIDN:CAB03889.1;
A;Experimental source: clone Cl6D6
C;Genetics:

A;Gene: CESP:C16D6.1
A;Map position: X
A;Introns: 42/1

Query Match 81.2%; Score 39; DB 2; Length 173;
Best Local Similarity 85.7%; Pred. No. 63;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 RPLPPLP 10
||:||||
Db 121 RMPPLP 127

RESULT 32
A70856
probable lppZ protein - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: A70856
R;Cole, S.T.; Davies, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Bross, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: A70856
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-373 <COL>
A;Cross-references: UNIPROT:O53253; UNIPARC:UPI00000318FB; GB:AL021287; GB:AL123456; NID:
A;Experimental source: strain H37RV
C;Genetics:
A;Gene: lppZ

Query Match 81.2%; Score 39; DB 2; Length 373;
Best Local Similarity 75.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLXP 12
||||:|
Db 54 PLPPVPFP 61

RESULT 33
S60975
hypothetical protein YNL152w - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein N1765
C;Species: Saccharomyces cerevisiae
C;Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 09-Jul-2004
C;Accession: S60975; S63104; S63822
R;Nasr, F.; Becam, A.M.; Herbert, C.J.
submitted to the EMBL Data Library, October 1995
A;Description: The sequence of 36.8 kb from the left arm of chromosome XIV reveals 24 complete dystrophy kinase.
A;Reference number: S60958
A;Accession: S60975
A;Molecule type: DNA
A;Residues: 1-409 <NAS>
A;Cross-references: UNIPROT:P53901; UNIPARC:UPI000013BB88; EMBL:X92517; NID:g1050783; PID:g1050783
R;Nasr, F.; Becam, A.M.; Herbert, C.
submitted to the Protein Sequence Database, April 1996
A;Reference number: S62967
A;Accession: S63104
A;Molecule type: DNA
A;Residues: 1-409 <NAW>
A;Cross-references: UNIPARC:UPI000013BB88; EMBL:271428; NID:g1302109; PID:e239813; PID:g1302109
A;Experimental source: strain S288C
R;Nasr, F.; Becam, A.M.; Herbert, C.J.
Yeast 12, 169-175, 1996
A;Title: The sequence of 36.8 kb from the left arm of chromosome XIV reveals 24 complete dystrophy kinase.
A;Reference number: S63805; MUID:96287653; PMID:8686380

A;Accession: S63822
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-409 <NAF>
A;Cross-references: UNIPARC:UPI000013BB88; EMBL:X92517; NID:g1050783; PID:CAA63287.1; PID:CAA63287.1; P
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995
C;Genetics:
A;Cross-references: SGD:S0005096
A;Map position: 14L
A;Note: YNL152w

Query Match 81.2%; Score 39; DB 2; Length 409;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 RPLPPLP 10
||:||||
Db 159 RPIPLP 165

RESULT 34
A34794
B-cell CLL/lymphoma 3 (BCL3) protein - human
C;Species: Homo sapiens (man)
C;Date: 13-Jul-1990 #sequence_revision 13-Jul-1990 #text_change 31-Dec-2004
C;Accession: A34794
R;Ohno, H.; Takimoto, G.; McKeithan, T.W.
Cell 60, 991-997, 1990
A;Title: The candidate proto-oncogene bcl-3 is related to genes implicated in cell linear
A;Reference number: A34794; MUID:90199880; PMID:2180580
A;Accession: A34794
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-446 <OHN>
A;Cross-references: UNIPROT:P20749; UNIPARC:UPI000012682B; GB:M31732; NID:g179375; PIDN:
C;Genetics:
A;Gene: GDB:BCL3; BCL4; D19S37
A;Cross-references: GDB:120561; OMIM:109560
A;Map position: 19q13.2-19q13.2
C;Superfamily: EGF homology
C;Keywords: transcription regulation

Query Match 81.2%; Score 39; DB 2; Length 446;
Best Local Similarity 66.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
||:||||
Db 435 RPVPSPAP 443

RESULT 35
A75464
probable zinc metal1 proteinase - Deinococcus radiodurans (strain. R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: A75464
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: A75464
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-472 <WHI>
A;Cross-references: UNIPROT:Q9RVZ5; UNIPARC:UPI000000D3D6E; GB:AE001941; GB:AE000513; NID:
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR0875
A;Map position: 1

Query Match 81.2%; Score 39; DB 2; Length 472;
 Best Local Similarity 77.8%; Pred. No. 1.8e+02;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
 | | | | | | |
 Db 459 RPLPPLPEP 467

RESULT 36
 C49507
 potassium channel Kv1.5, form 3 - mouse
 C;Species: Mus musculus (house mouse)
 C;Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 17-Nov-2000
 C;Accession: C49507
 R;Attali, B.; Lesage, F.; Ziliani, P.; Guillemare, E.; Honore, E.; Waldmann, R.; Hugnot, J. Biol. Chem. 268, 24283-24289, 1993
 A;Title: Multiple mRNA isoforms encoding the mouse cardiac Kv1-5 delayed rectifier K(+) A;Reference number: A49507; MUID:94043264; PMID:8226976
 A;Accession: C49507
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-514 <ATT>
 A;Cross-references: UNIPARC:UPI00001779F5; GB:L22218
 C;Superfamily: potassium channel protein drk1
 C;Keywords: alternative splicing

Query Match 81.2%; Score 39; DB 2; Length 514;
 Best Local Similarity 85.7%; Pred. No. 2e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 RPLPPLP 10
 | | | | | | |
 Db 65 RPLPPMP 71

RESULT 37
 A49507
 potassium channel Kv1.5 - mouse
 C;Species: Mus musculus (house mouse)
 C;Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
 C;Accession: A49507; B49507
 R;Attali, B.; Lesage, F.; Ziliani, P.; Guillemare, E.; Honore, E.; Waldmann, R.; Hugnot, J. Biol. Chem. 268, 24283-24289, 1993
 A;Title: Multiple mRNA isoforms encoding the mouse cardiac Kv1-5 delayed rectifier K(+) A;Reference number: A49507; MUID:94043264; PMID:8226976
 A;Accession: A49507
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-602 <ATT>
 A;Cross-references: UNIPROT:Q61762; UNIPARC:UPI0000028EB9; GB:L22218; NID:G435603; PIDN: A49507
 A;Accession: B49507
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 201-602 <AT2>
 A;Cross-references: UNIPARC:UPI000002A65A; GB:L22218
 C;Superfamily: potassium channel protein drk1
 C;Keywords: alternative splicing

Query Match 81.2%; Score 39; DB 2; Length 602;
 Best Local Similarity 85.7%; Pred. No. 2.3e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 RPLPPLP 10
 | | | | | | |
 Db 65 RPLPPMP 71

RESULT 38
 T49494
 condensin complex component cnd2 related protein [imported] - Neurospora crassa
 N;Alternate names: protein B14D6.460
 C;Species: Neurospora crassa

C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-May-2004
 C;Accession: T49494
 R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura, submitted to the Protein Sequence Database, May 2000
 A;Reference number: Z25022
 A;Accession: T49494
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-832 <SCH>
 A;Cross-references: UNIPARC:UPI00001798B1; EMBL:AL356173; GSPDB:GN00116; NCSP:B14D6.460
 A;Experimental source: BAC clone B14D6; strain OR74A
 C;Genetics:
 A;Gene: NCSP:B14D6.460
 A;Map position: 6
 A;Introns: 122/3
 C;Superfamily: chromosome condensation complex condensin, subunit H

Query Match 81.2%; Score 39; DB 2; Length 832;
 Best Local Similarity 77.8%; Pred. No. 3.2e+02;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
 | | | | | | |
 Db 754 RRLPPLPTP 762

RESULT 39
 T31742
 hypothetical protein C05C8.4 - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
 C;Accession: T31742
 R;Sammons, L.; Wohldmann, P.
 submitted to the EMBL Data Library, July 1997
 A;Description: The sequence of C. elegans cosmid C05C8.
 A;Reference number: Z21078
 A;Accession: T31742
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-1106 <SAM>
 A;Cross-references: UNIPROT:O16310; UNIPARC:UPI00000749AE; EMBL:AF016430; PIDN:AAB65371.1
 A;Experimental source: strain Bristol N2; clone C05C8
 C;Genetics:
 A;Gene: CESP:C05C8.4
 A;Map position: 5
 A;Introns: 25/3; 78/3; 117/1; 245/1; 591/1; 787/1; 1008/2

Query Match 81.2%; Score 39; DB 2; Length 1106;
 Best Local Similarity 66.7%; Pred. No. 4.4e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
 | | | | | | |
 Db 1038 RMPPPRMP 1046

RESULT 40
 T42230
 AF4 protein - mouse (fragment)
 N;Alternate names: serine/proline-rich FEL protein
 C;Species: Mus musculus (house mouse)
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C;Accession: T42230
 R;Baskaran, K.; Erfurth, F.; Taborn, G.; Copeland, N.G.; Gilbert, J.; Jenkins, N.A.; Ian submitted to the EMBL Data Library, July 1997
 A;Description: Cloning and developmental expression of the murine homolog of the acute le A;Reference number: Z22090
 A;Accession: T42230
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: mRNA
 A;Residues: 1-1211 <BAS>
 A;Cross-references: UNIPROT:O35233; UNIPARC:UPI0000029466; EMBL:AF013131; NID:G2582018; A;Experimental source: strain CD1

C;Genetics:
A;Gene: Af4
A;Map position: 5

Query Match 81.2%; Score 39; DB 2; Length 1211;
Best Local Similarity 66.7%; Pred. No. 4.8e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
|||:|
Db 655 RPKPVPPTP 663

RESULT 41
T42625
AF-4 protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C;Accession: T42625
R;Isnard, P.; Depetris, D.; Mattei, M.G.; Ferrier, P.; Djabali, M.
Mamm. Genome 9, 1065-1068, 1998
A;Title: CDNA cloning, expression and chromosomal localization of the murine AF-4 gene
A;Reference number: Z22176; MUID:99099257; PMID:9880680
A;Accession: T42625
A;Status: preliminary; translated from GB/EMBL/DBJB
A;Molecule type: mRNA
A;Residues: 1-1217 <ISN>
A;Cross-references: UNIPROT:O88573; UNIPARC:UPI0000029921; EMBL:AF074266; NID:g33328189;
A;Experimental source: strain M; thymus

Query Match 81.2%; Score 39; DB 2; Length 1217;
Best Local Similarity 66.7%; Pred. No. 4.8e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
|||:|
Db 660 RPKPVPPTP 668

RESULT 42
B45344
probable nuclear antigen - suid herpesvirus 1 (strain Kaplan)
C;Species: suid herpesvirus 1
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C;Accession: B45344
R;Vlcek, C.; Kozmik, Z.; Paces, V.; Schirm, S.; Schwyzer, M.
Virology 179, 365-377, 1990
A;Title: Pseudorabies virus immediate-early gene overlaps with an oppositely oriented or
A;Reference number: A45344; MUID:91021039; PMID:2171211
A;Accession: B45344
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-1733 <VLC>
A;Cross-references: UNIPROT:P33485; UNIPARC:UPI0000138ADE; GB:M34651; NID:g3334070; PIDN:
C;Superfamily: pseudorabies virus 1 nuclear antigen

Query Match 81.2%; Score 39; DB 1; Length 1733;
Best Local Similarity 77.8%; Pred. No. 7e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
|||:|
Db 279 RPPPLPPP 287

RESULT 43
B40505
hypothetical protein - suid herpesvirus 1 (strain Indiana-Funkhuser or Becker)
C;Species: suid herpesvirus 1
C;Date: 10-Apr-1992 #sequence_revision 10-Apr-1992 #text_change 09-Jul-2004
C;Accession: B40505
R;Cheung, A.K.
J. Virol. 65, 5260-5271, 1991

A;Title: Cloning of the latency gene and the early protein 0 gene of pseudorabies virus.
A;Reference number: A40505; MUID:91374576; PMID:1654441
A;Accession: B40505
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1958 <CHE>
A;Cross-references: UNIPROT:Q69340; UNIPARC:UPI00000F2308; GB:M57505; NID:g334066; PIDN:
C;Superfamily: pseudorabies virus 1 nuclear antigen

Query Match 81.2%; Score 39; DB 2; Length 1958;
Best Local Similarity 77.8%; Pred. No. 7.9e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
|||:|
Db 488 RPPPLPPP 496

RESULT 44
S72349
nonstructural polyprotein - eastern equine encephalomyelitis virus
N;Alternate names: nonstructural protein NSP1; nonstructural protein NSP2; nonstructural
C;Species: eastern equine encephalomyelitis virus
C;Date: 04-May-1998 #sequence_revision 15-May-1998 #text_change 09-Jul-2004
C;Accession: S72349
R;Weaver, S.C.; Hagenbaugh, A.; Bellew, L.A.; Netesov, S.V.; Volchkov, V.E.; Chang, G.J.;
Virology 197, 375-390, 1993
A;Title: A comparison of the nucleotide sequences of eastern and western equine encephal
A;Reference number: S72349; MUID:94025587; PMID:8105605
A;Accession: S72349
A;Status: preliminary
A;Molecule type: genomic RNA
A;Residues: 1-2493 <WEA>
A;Cross-references: UNIPROT:Q88789; UNIPARC:UPI00000F4983; EMBL:U01034; NID:g393006; PID:
A;Note: readthrough of the terminator UGA occurs between the codons AAU for residue 1878
A;Note: the readthrough stopcodon UGA for residue 1879 is translated as X
C;Superfamily: Semliki Forest virus nonstructural protein

Query Match 81.2%; Score 39; DB 2; Length 2493;
Best Local Similarity 66.7%; Pred. No. 1e+03;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
|||:|
Db 1670 RPAPVPVP 1678

RESULT 45
S26372
nonstructural polyprotein - eastern equine encephalomyelitis virus
N;Contains: nonstructural protein NSP1; nonstructural protein NSP2; nonstructural protei
C;Species: eastern equine encephalomyelitis virus
C;Date: 06-Jan-1994 #sequence_revision 24-Jul-1997 #text_change 31-Dec-2004
C;Accession: S26369; S26372
R;Volchkov, V.E.; Volchkova, V.A.; Netesov, S.V.
Mol. Gen. Mikrobiol. Virusol. 5, 8-15, 1991
A;Title: Complete nucleotide sequence of the eastern equine encephalomyelitis virus geno
A;Reference number: S26369; MUID:91375524; PMID:1896061
A;Accession: S26369
A;Molecule type: genomic RNA
A;Residues: 1-2493 <VOL>
A;Cross-references: UNIPROT:Q66581; UNIPARC:UPI0000178569; EMBL:X63135
A;Note: sequence could not be checked because of bad print in paper
A;Note: readthrough of the terminator UGA occurs between the codons AAU for residue 1878
A;Note: the readthrough stopcodon UGA for residue 1879 is translated as X
A;Accession: S26372
A;Molecule type: mRNA
A;Residues: 1886-2493 <VO2>
A;Cross-references: UNIPARC:UPI00000F0DC1; EMBL:X63135; NID:g59185; PIDN:CAA44847.1; PID
A;Note: sequence could not be checked because of bad print in paper
A;Note: this reading frame extends between two stop codons and does not begin with a sta
C;Keywords: nonstructural protein; polyprotein
F;1-532/Product: nonstructural protein NSP1 #status predicted <NS1>

F;533-1326/Product: nonstructural protein NSP2 #status predicted <NS2>
F;1327-1878/Product: nonstructural protein NSP3 #status predicted <NS3>
F;1886-2493/Product: nonstructural protein NSP4 #status predicted <NS4>

Query Match 81.2%; Score 39; DB 2; Length 2493;
Best Local Similarity 66.7%; Pred. No. 1e+03;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RPLPPLXP 12
|||:|
Db 1670 RPAPPVPV 1678

RESULT 46
E87601
OmpA family protein [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: E87601
R;Niernman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolo
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: E87601
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-191 <STO>
A;Cross-references: UNIPROT:Q9A4I6; UNIPARC:UPI00000C7879; GB:AE005673; NID:gl3424457; E
C;Genetics:
A;Gene: CC2845

Query Match 79.2%; Score 38; DB 2; Length 191;
Best Local Similarity 66.7%; Pred. No. 96;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RPLPPLXP 12
:|||||
Db 62 QPLPPAPLP 70

RESULT 47
T04353
DNA binding protein - rice
C;Species: Oryza sativa (rice)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T04353
R;Hsing, Y.C.; Tsao, C.V.; Chow, T.; Hsieh, J.; Chen, Z.
submitted to the EMBL Data Library, April 1995
A;Description: Rice early embryogenesis gene.
A;Reference number: Z14889
A;Accession: T04353
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-217 <HSI>
A;Cross-references: UNIPROT:P93426; UNIPARC:UPI000009C6C6; EMBL:U25283; NID:gl753084; PI
A;Experimental source: cv. Tainung 67, seed

Query Match 79.2%; Score 38; DB 2; Length 217;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RPLPPLXP 12
||:|||||
Db 13 RPMSPLPAP 21

RESULT 48
C75539
conserved hypothetical protein - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C;Accession: C75539
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma'
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: C75539
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-222 <WHI>
A;Cross-references: UNIPROT:Q9RXN1; UNIPARC:UPI00000C171C; GB:AE001889; GB:AE000513; NID
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR0279
A;Map position: 1

Query Match 79.2%; Score 38; DB 2; Length 222;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RPLPPLP 10
||:||||
Db 201 RPVPPLP 207

RESULT 49
A38582
pollen allergen p1b precursor - perennial ryegrass
N;Alternate names: 30K allergen
C;Species: Lolium perenne (perennial ryegrass)
C;Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 07-May-1999
C;Accession: A38582; S38290
R;Singh, M.B.; Hough, T.; Theerakulpisut, P.; Avjioğlu, A.; Davies, S.; Smith, P.M.; Tay
Proc. Natl. Acad. Sci. U.S.A. 88, 1384-1388, 1991
A;Title: Isolation of cDNA encoding a newly identified major allergenic protein of rye-gr
A;Reference number: A38582; MUID:91142177; PMID:1671715
A;Accession: A38582
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-308 <SIN>
A;Cross-references: UNIPARC:UPI0000177EPF; GB:M59163
R;Petersen, A.; Schramm, G.; Becker, W.M.; Schlaak, M.
Biol. Chem. Hoppe-Seyler 374, 855-861, 1993
A;Title: Comparison of four grass pollen species concerning their allergens of grass gro
A;Reference number: S38288; MUID:94092339; PMID:7505588
A;Accession: S38290
A;Molecule type: protein
A;Residues: 26-45 <PET>
A;Cross-references: UNIPARC:UPI00001474F3
C;Superfamily: grass pollen allergen IX
C;Keywords: pollen
F;1-25/Domain: signal sequence #status predicted <SIG>
F;26-308/Product: pollen allergen p1b #status experimental <MAT>

Query Match 79.2%; Score 38; DB 2; Length 308;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RPLPPLXP 12
:|:|||||
Db 70 QPWPPLPTP 78

RESULT 50
A31978
ADP,ATP carrier protein AAC2 precursor - yeast (Saccharomyces cerevisiae)
N;Alternate names: adenine nucleotide carrier; protein YBL030C; protein YBL0421
C;Species: Saccharomyces cerevisiae
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: A31978; S36419; S45764; S50304; S42759
R;Lawson, J.E.; Douglas, M.G.
J. Biol. Chem. 263, 14812-14818, 1988

A;Title: Separate genes encode functionally equivalent ADP/ATP carrier proteins in Sacch
A;Reference number: A31978; MUID:89008354; PMID:2844786
A;Accession: A31978
A;Molecule type: DNA
A;Residues: 1-318 <LAW>
A;Cross-references: UNIPROT:P18239; UNIPARC:UPI000012561E; EMBL:J04021; NID:g170957; PID
R;Trezeguet, V.; le Saux, A.; Lauquin, J.M.
submitted to the EMBL Data Library, August 1993
A;Reference number: S36419
A;Accession: S36419
A;Molecule type: DNA
A;Residues: 1-318 <TRE>
A;Cross-references: UNIPARC:UPI000012561E; EMBL:X74427; NID:g396554; PIDN:CAA52446.1; PI
R;Goffeau, A.; Jonniaux, J.L.; Purnelle, B.; Skala, J.; de Wergifosse, P.; van Dyck, L.
submitted to the Protein Sequence Database, August 1994
A;Reference number: S45745
A;Accession: S45764
A;Molecule type: DNA
A;Residues: 1-318 <GOF>
A;Cross-references: UNIPARC:UPI000012561E; EMBL:X35791; NID:g536034; PIDN:CAA84850.1; PI
R;van Dyck, L.; Jonniaux, J.L.; de Melo Barreiros, T.; Kleine, K.; Goffeau, A.
Yeast 10, 1663-1673, 1994
A;Title: Analysis of a 17.4 kb DNA segment of yeast chromosome II encompassing the ribos
human proliferation-associated p120 antigen.
A;Reference number: S50299; MUID:95242843; PMID:7725803
A;Accession: S50304
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-318 <VAN>
A;Cross-references: UNIPARC:UPI000012561E; EMBL:X77291; NID:g602888; PIDN:CAA54501.1; PI
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1994
R;Kolarov, J.; Kolarova, N.; Nelson, N.
J. Biol. Chem. 265, 12711-12716, 1990
A;Title: A third ADP/ATP translocator gene in yeast.
A;Reference number: A36582; MUID:90324269; PMID:2165073
A;Accession: S42759
A;Molecule type: DNA
A;Residues: 1-57, 'I', 59-64, 'K', 66-67, 'S', 69-70, 'V', 72-78, 'K', 80-82, 'L', 84-112, 'L', 114-12
A;Cross-references: UNIPARC:UPI0000168AD7; EMBL:M34075; NID:g170961; PIDN:AAA97484.1; PI
C;Genetics:
A;Gene: SGD:PET9; AAC2; MIPS:YBL030c
A;Cross-references: MIPS:YBL030c; SGD:S0000126
A;Map position: 2L
A;Genome: nuclear
C;Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology
C;Keywords: duplication; homodimer; mitochondrial inner membrane; mitochondrion; transme
F;1-20/Domain: transit peptide (mitochondrion) #status predicted <TNP>
F;20-115/Domain: ADP,ATP carrier protein repeat homology <ACP1>
F;21-318/Product: ADP,ATP carrier protein AAC2 #status predicted <MAT>
F;21-37/Domain: transmembrane #status predicted <TM1>
F;124-218/Domain: ADP,ATP carrier protein repeat homology <ACP2>
F;131-147/Domain: transmembrane #status predicted <TM2>
F;186-202/Domain: transmembrane #status predicted <TM3>
F;217-233/Domain: transmembrane #status predicted <TM4>
F;224-312/Domain: ADP,ATP carrier protein repeat homology <ACP3>

Query Match 79.2%; Score 38; DB 1; Length 318;
Best Local Similarity 75.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 PLPPLXP 12
| | | | |
Db 10 PLPPAPAP 17

Search completed: April 6, 2006, 09:39:41
Job time : 48.7895 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2006, 09:23:24 ; Search time 145.895 Seconds
(without alignments)
58.030 Million cell updates/sec

Title: US-10-632-388-307
Perfect score: 48
Sequence: 1 XXXRPLPLXP 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : Uniprot 05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	100.0	1112	2	Q4I257_GIBZEA
2	47	97.9	613	1	KCNAS5_HUMAN
3	47	97.9	613	2	Q4VAJ1_HUMAN
4	46	95.8	282	1	TNLF6_FIG
5	46	95.8	1794	2	Q55ZA2_CRYNE
6	46	95.8	1811	2	Q5XNL7_CRYNE
7	45	93.8	85	1	FDSCP_HUMAN
8	45	93.8	85	2	Q540F3_HUMAN
9	45	93.8	519	2	Q91NH0_9PARA
10	44	91.7	147	2	Q6ZS83_HUMAN
11	44	91.7	252	2	Q54M51_DICDI
12	44	91.7	355	2	Q5M9F4_RAT
13	43	89.6	162	2	Q43516_LYCERS
14	43	89.6	162	2	Q6KA56_ORYSA
15	43	89.6	165	2	Q8BXY0_MOUSE
16	43	89.6	170	2	Q8H5L8_ORYSA
17	43	89.6	209	1	CXX1_HUMAN
18	43	89.6	209	2	Q6IBF1_HUMAN
19	43	89.6	238	2	Q51CQ5_ENTHI
20	43	89.6	243	2	Q4QHF1_LEIMA
21	43	89.6	287	1	UDU3_ARATH
22	43	89.6	462	2	Q9FB15_9ACTO
23	43	89.6	486	2	Q9KW38_9RICK
24	43	89.6	486	2	Q73IY6_WOLPM
25	43	89.6	487	2	Q63IB6_BURPS
26	43	89.6	488	2	Q4P068_USTMA
27	43	89.6	493	2	Q52SK2_9RICK
28	43	89.6	551	2	Q4WXJ5_ASPFU
29	43	89.6	554	2	Q6Z7E0_ORYSA
30	43	89.6	640	2	Q55P29_CRYNE
31	43	89.6	745	2	Q6Z0N5_ORYSA
					Q4I257_gibberella
					P22460_homo sapien
					Q4vaj1_homo sapien
					Q9bea8_sus scrofa
					Q55za2_cryptococcu
					Q5kn17_cryptococcu
					Q8nfu4_homo sapien
					Q540f3_homo sapien
					Q91nh0_tioaman viru
					Q6zs83_homo sapien
					Q54m51_dictyosteli
					Q5m9f4_rattus norv
					Q43516_lycopersico
					Q6ka56_oryza sativ
					Q8bxy0_mus musculu
					Q8h5l8_oryza sativ
					O15255_homo sapien
					Q6ibf1_homo sapien
					Q5lccq5_entamoeba h
					Q4ghf1_leishmania
					Q9fhd3_arabidopsis
					Q9fb15_streptomyce
					Q9kw38_wolbachia s
					Q73iy6_wolbachia p
					Q63ib6_burkholderi
					Q4p068_ustilago ma
					Q52sk2_wolbachia e
					Q4wxj5_aspergillus
					Q6z7e0_oryza sativ
					Q55p29_cryptococcu
					Q6z0n5_oryza sativ

32	43	89.6	851	2	Q4T7J1_TETNG	Q4t7j1 tetraodon n
33	43	89.6	963	2	Q4SDM9_TETNG	Q4sdm9 tetraodon n
34	43	89.6	969	2	Q7RYU1_NEUCR	Q7ryu1 neurospora
35	43	89.6	1299	2	Q9U5X0_DROME	Q9u5x0 drosophila
36	43	89.6	1299	2	Q9V4I9_DROME	Q9v4i9 drosophila
37	43	89.6	2145	2	Q4QEP3_LEIMA	Q4qep3 leishmania
38	43	89.6	2432	2	Q5XMJ3_9PICO	Q5xmj3 aichi virus
39	43	89.6	2432	2	Q91QP4_9PICO	Q91qp4 aichi virus
40	43	89.6	2433	2	Q91464_9PICO	Q91464 aichi virus
41	43	89.6	2433	2	Q4TWI0_9PICO	Q4twi0 aichi virus
42	42	87.5	80	2	Q5WLC4_BACSK	Q5wlc4 bacillus cl
43	42	87.5	106	2	Q4H8A2_9DEIO	Q4h8a2 deinococcus
44	42	87.5	134	2	Q5AGI6_CANAL	Q5agi6 candida alb
45	42	87.5	144	2	Q5Z6V2_ORYSA	Q5zbv2 oryza sativ
46	42	87.5	150	2	Q5ZB26_ORYSA	Q5zb26 oryza sativ
47	42	87.5	151	2	Q8BGM6_MOUSE	Q8bgm6 m mus muscu
48	42	87.5	176	2	Q8PNK3_XANAC	Q8pnk3 xanthomonas
49	42	87.5	188	2	Q9D7T0_MOUSE	Q9d7t0 mus musculu
50	42	87.5	198	2	Q01621_CABEL	Q01621 caenorhabdi
51	42	87.5	280	1	TNLF6_CERTO	Q9bdn1 c tumor nec
52	42	87.5	280	1	TNLF6_MACFA	P63308 m tumor nec
53	42	87.5	280	1	TNLF6_MACMU	P63307 m tumor nec
54	42	87.5	280	1	TNLF6_MACNE	P63306 m tumor nec
55	42	87.5	280	2	Q861W5_FELCA	Q861w5 felis silve
56	42	87.5	281	1	TNLF6_HUMAN	P48023 h tumor nec
57	42	87.5	281	2	Q53ZZ1_HUMAN	Q53zz1 homo sapien
58	42	87.5	298	2	Q93F75_9BACI	Q93f75 bacillus sp
59	42	87.5	313	2	Q60DI8_ORYSA	Q60di8 oryza sativ
60	42	87.5	355	2	Q84YW6_ORYSA	Q84yw6 oryza sativ
61	42	87.5	473	2	Q39620_CHLRE	Q39620 chlamydomon
62	42	87.5	503	2	Q9KW45_9RICK	Q9kw45 wolbachia s
63	42	87.5	519	2	Q7UII0_RHOBA	Q7uii0 rhodopirell
64	42	87.5	584	2	Q4RH23_TETNG	Q4rh23 tetraodon n
65	42	87.5	594	2	Q84Z97_ORYSA	Q84z97 oryza sativ
66	42	87.5	606	2	Q852P0_VOLCA	Q852p0 volvox cart
67	42	87.5	643	2	Q8CB81_MOUSE	Q8cb81 mus musculu
68	42	87.5	665	2	Q80TL0_MOUSE	Q80tl0 mus musculu
69	42	87.5	749	2	Q5SX30_MOUSE	Q5sx30 mus musculu
70	42	87.5	750	2	Q80Z30_RAT	Q80z30 rattus norv
71	42	87.5	757	2	Q7LAF3_HUMAN	Q7laf3 homo sapien
72	42	87.5	759	2	Q9UPT0_HUMAN	Q9upt0 homo sapien
73	42	87.5	766	2	Q8WY54_HUMAN	Q8wy54 homo sapien
74	42	87.5	773	2	Q4SHR4_TETNG	Q4shr4 tetraodon n
75	42	87.5	837	2	Q5QMC9_ORYSA	Q5qmc9 oryza sativ
76	42	87.5	853	2	Q14333_HUMAN	Q14333 homo sapien
77	42	87.5	885	2	Q4RWK8_TETNG	Q4rwk8 tetraodon n
78	42	87.5	1035	2	Q5RIG8_BRABE	Q5rig8 brachydanio
79	42	87.5	1077	2	Q5LCE8_BACFN	Q5lce8 bacteroides
80	42	87.5	1077	2	Q64TJ3_BACFR	Q64tj3 bacteroides
81	42	87.5	1157	2	Q872I7_NEUCR	Q872i7 neurospora
82	42	87.5	1272	2	Q54PT8_DICDI	Q54pt8 dictyosteli
83	42	87.5	1430	1	CSKI1_RAT	Q8vbk2 rattus norv
84	42	87.5	2775	2	Q5B770_EMENI	Q5b770 aspergillus
85	41	85.4	40	2	Q62KF6_BURMA	Q62kf6 burkholderi
86	41	85.4	101	2	Q8QS70_9BETA	Q8qs70 pongine her
87	41	85.4	128	2	Q8STT3_ENCCU	Q8stt3 encephalito
88	41	85.4	128	2	Q67TS6_ORYSA	Q67ts6 oryza sativ
89	41	85.4	130	2	Q6YUJ6_ORYSA	Q6yuj6 oryza sativ
90	41	85.4	131	2	Q4URI2_XANCP	Q4uri2 xanthomonas
91	41	85.4	131	2	Q8PC16_XANCP	Q8pci16 xanthomonas
92	41	85.4	151	2	Q9FI95_ARATH	Q9fi95 arabidopsis
93	41	85.4	156	1	VPG_PLRV	P10471 potato leaf
94	41	85.4	156	2	Q84828_PLRV	Q84828 potato leaf
95	41	85.4	156	2	Q8QYN5_PLRV	Q8qyn5 potato leaf
96	41	85.4	156	2	Q8QYQ5_PLRV	Q8qyq5 potato leaf
97	41	85.4	156	2	Q99GS2_PLRV	Q99gs2 potato leaf
98	41	85.4	168	2	Q67TQ3_ORYSA	Q67tq3 oryza sativ
99	41	85.4	199	2	Q7XR59_ORYSA	Q7xr59 oryza sativ
100	41	85.4	240	2	Q4HY80_GIBZE	Q4hy80 gibberella
101	41	85.4	260	2	Q82HP0_STRAW	Q82hp0 streptomyce
102	41	85.4	271	2	Q80WJ0_MOUSE	Q80wj0 mus musculu
103	41	85.4	279	2	Q6N908_RHOBA	Q6n908 rhodopseudo
104	41	85.4	294	2	Q8W0Q8_SORBI	Q8w0q8 sorghum bic

105	41	85.4	320	2	Q6YY27_ORYSA	Q6yy27	oryza sativ
106	41	85.4	321	2	Q9HCN7_HUMAN	Q9hcn7	homo sapien
107	41	85.4	337	2	Q5FPY8_GLUOX	Q5fpy8	gluconobact
108	41	85.4	339	2	Q9UIF2_HUMAN	Q9uif2	homo sapien
109	41	85.4	354	2	Q4I2G8_GIBZE	Q4i2g8	gibberella
110	41	85.4	355	2	Q4ZVK0_PSESY	Q4zvk0	pseudomonas
111	41	85.4	378	2	Q5YME9_NOCFA	Q5yme9	nocardia fa
112	41	85.4	386	2	O17174_CAEEL	O17174	caenorhabdi
113	41	85.4	401	2	Q6PA50_XENLA	Q6pa50	xenopus lae
114	41	85.4	404	2	Q6DF84_XENLA	Q6df84	xenopus lae
115	41	85.4	423	2	Q872X2_NEUCR	Q872x2	neurospora
116	41	85.4	431	2	Q8IRH4_DROME	Q8irh4	drosophila
117	41	85.4	443	2	Q67VV7_ORYSA	Q67vv7	oryza sativ
118	41	85.4	459	1	MIG6_RAT	P05432	rattus norv
119	41	85.4	461	1	MIG6_MOUSE	Q99jz7	mus musculu
120	41	85.4	462	1	MIG6_HUMAN	Q9ujm3	homo sapien
121	41	85.4	477	2	Q9NSQ8_HUMAN	Q9nsq8	homo sapien
122	41	85.4	488	2	Q51A29_ENTHI	Q51a29	entamoeba h
123	41	85.4	495	2	Q9H5I7_HUMAN	Q9h5i7	homo sapien
124	41	85.4	499	2	Q51A34_ENTHI	Q51a34	entamoeba h
125	41	85.4	526	2	Q501W1_RAT	Q501w1	rattus norv
126	41	85.4	545	1	FGR_FSVGR	P00544	feline sarc
127	41	85.4	548	1	PTP61_DROME	Q9w0g1	drosophila
128	41	85.4	566	1	UBP21_MOUSE	Q9qz16	mus musculu
129	41	85.4	586	2	Q9DSX3_MOUSE	Q9d5x3	mus musculu
130	41	85.4	601	1	KCNA5_MUSPF	P79197	mustela put
131	41	85.4	627	2	Q9XVH7_CAEEL	Q9xvh7	caenorhabdi
132	41	85.4	628	2	Q4NYQ7_9DELT	Q4nyq7	anaeromyxob
133	41	85.4	631	2	Q4XOH0_ASPFU	Q4xoh0	aspergillus
134	41	85.4	631	2	Q5XIL5_RAT	Q5xil5	rattus norv
135	41	85.4	641	1	MA1A2_HUMAN	O60476	homo sapien
136	41	85.4	658	2	Q66HC8_RAT	Q66hc8	rattus norv
137	41	85.4	664	2	Q4RTR5_TETNG	Q4trr5	tetraodon n
138	41	85.4	673	2	Q5EBP4_MOUSE	Q5ebp4	mus musculu
139	41	85.4	675	2	Q80WJ1_MOUSE	Q80wj1	mus musculu
140	41	85.4	678	2	Q7RWV5_NEUCR	Q7rwv5	neurospora
141	41	85.4	678	2	Q4IIQ6_GIBZE	Q4ilq6	gibberella
142	41	85.4	740	2	Q59U10_CANAL	Q59u10	candida alb
143	41	85.4	742	2	Q59U34_CANAL	Q59u34	candida alb
144	41	85.4	753	2	Q6PIJ3_HUMAN	Q6pij3	homo sapien
145	41	85.4	760	2	Q8NBU7_HUMAN	Q8nbu7	homo sapien
146	41	85.4	774	1	GLIS3_HUMAN	Q8nea6	homo sapien
147	41	85.4	775	2	Q5VZV9_HUMAN	Q5vzv9	homo sapien
148	41	85.4	775	2	Q4SME7_TETNG	Q4sme7	tetraodon n
149	41	85.4	780	2	Q9LK78_ARATH	Q9lk78	arabidopsais
150	41	85.4	781	2	Q823Y3_CHLCV	Q823y3	chlamydophi
151	41	85.4	782	2	Q8GX37_ARATH	Q8gx37	arabidopsais
152	41	85.4	812	2	Q9DWG7_RCMVM	Q9dwg7	rat cytomeg
153	41	85.4	824	2	Q5KQN9_ORYSA	Q5kqn9	oryza sativ
154	41	85.4	847	2	Q6UXZ5_HUMAN	Q6uxz5	homo sapien
155	41	85.4	847	2	Q5R9H3_PONPY	Q5r9h3	pongo pygma
156	41	85.4	849	2	Q7T299_BRARE	Q7t299	brachydanio
157	41	85.4	874	2	Q4UFW5_THEAN	Q4ufw5	theileria a
158	41	85.4	894	2	O22824_ARATH	O22824	arabidopsais
159	41	85.4	934	2	Q8SZX4_DROME	Q8szx4	drosophila
160	41	85.4	1022	2	Q4RWV9_TETNG	Q4rww9	tetraodon n
161	41	85.4	1054	2	Q4Q0W2_LEIMA	Q4q0w2	leishmania
162	41	85.4	1125	2	Q5CHA3_CRYHO	Q5cha3	cryptospori
163	41	85.4	1151	2	Q4SEG7_TETNG	Q4seg7	tetraodon n
164	41	85.4	1204	2	Q8IXB2_HUMAN	Q8ixb2	homo sapien
165	41	85.4	1204	2	Q53EP0_HUMAN	Q53ep0	homo sapien
166	41	85.4	1204	2	Q5U5T8_HUMAN	Q5u5t8	homo sapien
167	41	85.4	1219	2	Q4S0T1_TETNG	Q4s0t1	tetraodon n
168	41	85.4	1241	2	Q9W3L1_DROME	Q9w3l1	drosophila
169	41	85.4	1292	2	Q20330_CAEEL	Q20330	caenorhabdi
170	41	85.4	1353	2	Q55UG1_CRYNE	Q55ug1	cryptococcu
171	41	85.4	1353	2	Q5KHV3_CRYNE	Q5khy3	cryptococcu
172	41	85.4	1359	2	Q7RXQ6_NEUCR	Q7rxq6	neurospora
173	41	85.4	1363	2	Q4STY2_TETNG	Q4sty2	tetraodon n
174	41	85.4	1370	2	Q7S0P7_NEUCR	Q7s0p7	neurospora
175	41	85.4	1379	2	Q5AZY0_EMENT	Q5azy0	aspergillus
176	41	85.4	1409	2	Q4X1X2_ASPFU	Q4x1x2	aspergillus
177	41	85.4	1473	2	Q4FVZ1_LEIMA	Q4fvz1	leishmania

41	85.4	1487	2	Q68E09_HUMAN	Q68e09	homo sapien
41	85.4	1600	2	Q4I646_GIBZE	Q4i646	gibberella
41	85.4	1651	2	Q6ZML7_HUMAN	Q6zml7	homo sapien
41	85.4	1837	2	Q9VCZ9_DROME	Q9vcz9	drosophila
41	85.4	2054	2	Q7QAT6_ANOGA	Q7qat6	anopheles g
40	83.3	111	1	Y3835_GLOVI	Q7nep3	gloeobacter
40	83.3	111	2	Q9WAZ9_9VIRU	Q9waz9	torque teno
40	83.3	111	2	Q9WB00_9VIRU	Q9wb00	torque teno
40	83.3	111	2	Q9WB03_9VIRU	Q9wb03	torque teno
40	83.3	112	2	Q9QU23_9VIRU	Q9qu23	torque teno
40	83.3	121	2	Q9DTD3_9VIRU	Q9dtd3	torque teno
40	83.3	128	2	Q8V7C9_9VIRU	Q8v7c9	torque teno
40	83.3	134	2	Q9X8L8_STRCO	Q9x8l8	streptomyce
40	83.3	156	1	VPG_PLRVI	P17524	potato leaf
40	83.3	156	1	VPG_PLRVR	P17523	potato leaf
40	83.3	156	1	VPG_PLRVW	P11625	potato leaf
40	83.3	156	2	Q77E88_PLRV	Q77e88	potato leaf
40	83.3	156	2	Q84839_PLRV	Q84839	potato leaf
40	83.3	156	2	Q8QYP2_PLRV	Q8qyp2	potato leaf
40	83.3	156	2	Q8QYP8_PLRV	Q8qyp8	potato leaf
40	83.3	156	2	Q8QYR2_PLRV	Q8qyr2	potato leaf
40	83.3	156	2	Q8QYR2_PLRV	Q8qyr2	potato leaf
40	83.3	156	2	Q99GS5_PLRV	Q99gs5	potato leaf
40	83.3	156	2	Q9YNY3_PLRV	Q9yny3	potato leaf
40	83.3	156	2	Q70805_9VIRU	Q70805	torque teno
40	83.3	156	2	Q9WSW8_9VIRU	Q9ws8	torque teno
40	83.3	157	2	Q9JH34_9VIRU	Q9jh34	torque teno
40	83.3	165	2	Q64M82_ORYSA	Q64m82	oryza sativ
40	83.3	172	2	Q9DS34_MOUSE	Q9d534	mus musculu
40	83.3	173	2	Q6ZOM9_ORYSA	Q6zom9	oryza sativ
40	83.3	204	2	Q9WSW6_9VIRU	Q9ws6	torque teno
40	83.3	205	2	Q655L3_ORYSA	Q655l3	oryza sativ
40	83.3	215	2	Q18661_CAEEL	Q18661	caenorhabdi
40	83.3	238	2	Q8GR34_DESVU	Q8gr34	desulfovibr
40	83.3	244	2	Q63NA8_BURPS	Q63na8	burkholderi
40	83.3	245	2	O01572_CABEL	O01572	caenorhabdi
40	83.3	247	2	Q4UG37_THEAN	Q4ug37	theileria a
40	83.3	247	2	Q4N9N7_THEPA	Q4n9n7	theileria p
40	83.3	273	2	O06170_MYCTU	O06170	mycobacteri
40	83.3	273	2	Q7TYF3_MYCBO	Q7tyf3	mycobacteri
40	83.3	275	2	Q7D707_MYCTU	Q7d707	mycobacteri
40	83.3	283	2	Q6ZJL2_ORYSA	Q6zjl2	oryza sativ
40	83.3	310	2	Q5YX54_NOCFA	Q5yx54	nocardia fa
40	83.3	382	2	Q60Q47_CAEBR	Q60q47	caenorhabdi
40	83.3	385	2	Q9GYM4_CAEEL	Q9gym4	caenorhabdi
40	83.3	385	2	Q6K649_ORYSA	Q6k649	oryza sativ
40	83.3	421	2	Q5SKB6_THET8	Q5skb6	thermus the
40	83.3	421	2	Q72KM8_THET2	Q72km8	thermus the
40	83.3	429	2	Q75M69_ORYSA	Q75m69	oryza sativ
40	83.3	488	2	Q5Z1J6_NOCFA	Q5z1j6	nocardia fa
40	83.3	501	2	Q5Z5Z3_ORYSA	Q5z5z3	oryza sativ
40	83.3	527	2	Q5JXK0_ORYSA	Q5jxk0	oryza sativ
40	83.3	530	2	Q5KC56_CRYNE	Q5kc56	cryptococcu
40	83.3	546	2	Q51MH6_MAGGR	Q5lmh6	magnaporthe
40	83.3	553	2	Q6DN78_LOLPR	Q6dn78	lolium pere
40	83.3	564	2	Q9LDA9_ARATH	Q9lda9	arabidopsais
40	83.3	566	2	O86854_STRCO	O86854	streptomyce
40	83.3	574	1	DPOA2_SCHPO	Q74946	schizosacch
40	83.3	574	2	Q59B22_ASPPA	Q69b22	aspergillus
40	83.3	577	2	Q75JH4_DICDI	Q75jh4	dictyosteli
40	83.3	601	2	Q4I5K8_GIBZE	Q4i5k8	gibberella
40	83.3	629	2	Q4FX88_LEIMA	Q4fx88	leishmania
40	83.3	748	2	Q51W59_MAGGR	Q51w59	magnaporthe
40	83.3	765	2	Q55UM0_CRYNE	Q55um0	cryptococcu
40	83.3	765	2	Q5KHS4_CRYNE	Q5khs4	cryptococcu
40	83.3	775	2	Q9JS20_CHLPN	Q9js20	chlamydia p
40	83.3	775	2	Q9Z879_CHLPN	Q9z879	chlamydia p
40	83.3	790	2	Q82LN3_STRAW	Q82ln3	streptomyce
40	83.3	791	1	SF3A1_MOUSE	Q8k4z5	mus musculu
40	83.3	791	2	Q5ZM84_CHICK	Q5zm84	gallus gall
40	83.3	793	1	SF3A1_HUMAN	Q15459	homo sapien
40	83.3	797	2	Q9RKS9_STRCO	Q9rks9	streptomyce
40	83.3	809	2	Q8RWQ1_ARATH	Q8rwq1	arabidopsais
40	83.3	809	2	Q8CA82_MOUSE	Q8ca82	mus musculu

251	40	83.3	864	1	ADA15_MOUSE	088839	mus musculus	324	39	81.2	675	2	Q6UEI2_PNECA	Q6uei2	pneumocysti
252	40	83.3	864	2	Q6P779_RAT	Q6p779	rattus norv	325	39	81.2	749	2	Q57VY5_9TRYP	Q57vy5	trypanosoma
253	40	83.3	929	2	Q59P09_CANAL	Q59p09	candida alb	326	39	81.2	753	2	Q6ZIP3_ORYSA	Q6zip3	oryza sativ
254	40	83.3	944	2	Q4Q9T9_LEIMA	Q4q9t9	leishmania	327	39	81.2	780	2	Q9K402_STRCO	Q9k402	streptomyce
255	40	83.3	1008	2	Q72KA4_THET2	Q72ka4	thermus the	328	39	81.2	788	2	Q69MT2_ORYSA	Q69mt2	oryza sativ
256	40	83.3	1008	2	Q5SJV3_THET8	Q5sjv3	thermus the	329	39	81.2	809	2	O76869_DROME	O76869	drosophila
257	40	83.3	1048	2	Q5AYK5_EMENI	Q5ayk5	aspergillus	330	39	81.2	824	2	Q4RID3_TETNG	Q4rid3	tetraodon n
258	40	83.3	1367	2	Q552E9_DICDI	Q552e9	dictyosteli	331	39	81.2	845	2	Q6ZH84_ORYSA	Q6zh84	oryza sativ
259	40	83.3	1424	2	Q5XJV6_MOUSE	Q5xjv6	mus musculus	332	39	81.2	851	1	DDX31_HUMAN	Q9h8h2	homo sapien
260	40	83.3	1424	2	Q52KF1_MOUSE	Q52kf1	mus musculus	333	39	81.2	853	2	Q4IAH9_GIBZE	Q4iah9	gibberella
261	40	83.3	2371	2	Q58NA5_CHLIN	Q58na5	chlamydomon	334	39	81.2	859	2	Q8IRV3_DROME	Q8irv3	drosophila
262	40	83.3	3889	2	Q6SSE8_CHLRE	Q6sse8	chlamydomon	335	39	81.2	861	2	Q55XX6_CRYNE	Q55xx6	cryptococcu
263	40	83.3	3930	2	Q54HJ6_DICDI	Q54hj6	dictyosteli	336	39	81.2	861	2	Q5KM19_CRYNE	Q5km19	cryptococcu
264	39	81.2	117	2	Q8RYA1_ORYSA	Q8rya1	oryza sativ	337	39	81.2	872	2	Q6ZA49_ORYSA	Q6za49	oryza sativ
265	39	81.2	127	2	Q95KG3_MACFA	Q95kg3	macaca fasc	338	39	81.2	887	2	Q9W4V1_DROME	Q9w4v1	drosophila
266	39	81.2	130	2	Q8WZJ3_9APHY	Q8wzj3	dictyonema	339	39	81.2	923	2	Q4I4L5_GIBZE	Q4i4l5	gibberella
267	39	81.2	130	2	Q5NN73_ZYMMO	Q5nn73	zymomonas m	340	39	81.2	939	2	Q4PG76_USTMA	Q4pg76	ustilago ma
268	39	81.2	132	2	Q57WK9_9TRYP	Q57wk9	trypanosoma	341	39	81.2	948	2	Q4TK59_9SPHN	Q4tk59	erythrobact
269	39	81.2	132	2	Q57YI6_9TRYP	Q57yi6	trypanosoma	342	39	81.2	962	2	Q61PF1_CAEBR	Q61pf1	caenorhabdi
270	39	81.2	134	2	Q9DUH9_9VIRU	Q9duh9	torque teno	343	39	81.2	1006	2	Q750H0_ASHGO	Q750h0	ashbya goss
271	39	81.2	137	2	Q8IYK6_HUMAN	Q8iyk6	homo sapien	344	39	81.2	1104	2	Q55HE0_CRYNE	Q55he0	cryptococcu
272	39	81.2	153	1	PRIMA_HUMAN	Q86xrs	homo sapien	345	39	81.2	1104	2	Q5K6X9_CRYNE	Q5k6x9	cryptococcu
273	39	81.2	153	1	PRIMA_MOUSE	Q810f0	mus musculus	346	39	81.2	1106	2	O16310_CAEEL	O16310	caenorhabdi
274	39	81.2	153	2	Q5TFG9_HUMAN	Q5tfg9	homo sapien	347	39	81.2	1141	2	Q7S421_NEUCR	Q7s421	neurospora
275	39	81.2	153	2	Q9I542_PSEAB	Q9i542	pseudomonas	348	39	81.2	1211	2	O35233_MOUSE	O35233	mus musculus
276	39	81.2	172	2	Q6Z4V3_ORYSA	Q6z4v3	oryza sativ	349	39	81.2	1217	1	APF1_MOUSE	O88573	mus musculus
277	39	81.2	173	2	Q62063_CAEEL	Q62063	caenorhabdi	350	39	81.2	1218	2	Q5XM34_MOUSE	O5xm34	mus musculus
278	39	81.2	189	2	Q84QR5_ORYSA	Q84qr5	oryza sativ	351	39	81.2	1233	2	Q7SCT1_NEUCR	Q7sct1	neurospora
279	39	81.2	216	2	Q4NPG8_9DELT	Q4npg8	anaeromyxob	352	39	81.2	1285	2	Q8K3T3_SPAJD	Q8k3t3	spalax juda
280	39	81.2	251	2	Q6IJW2_DROME	Q6ijw2	drosophila	353	39	81.2	1318	2	Q55LL3_CRYNE	Q55ll3	cryptococcu
281	39	81.2	252	2	Q5Z732_ORYSA	Q5z732	oryza sativ	354	39	81.2	1318	2	Q5KD53_CRYNE	Q5kd53	cryptococcu
282	39	81.2	307	2	Q9VZZ8_DROME	Q9vzz8	drosophila	355	39	81.2	1426	2	Q4RTA6_TETNG	Q4rta6	tetraodon n
283	39	81.2	309	2	Q8UZB4_9VIRU	Q8uzb4	grapevine f	356	39	81.2	1462	2	Q4TBB7_TETNG	Q4tbb7	tetraodon n
284	39	81.2	318	2	Q41545_WHEAT	Q41545	triticum ae	357	39	81.2	1698	2	Q4RZU0_TETNG	Q4rzu0	tetraodon n
285	39	81.2	323	2	Q7V622_PROMM	Q7v622	prochlororoc	358	39	81.2	1733	1	VNUA_PRVKA	P33485	pseudorabie
286	39	81.2	344	1	TILS_SYNEL	Q8dih2	synechococc	359	39	81.2	1733	2	Q5PPC8_9ALPH	Q5ppc8	suid herpes
287	39	81.2	355	1	SIRT6_HUMAN	Q8n6t7	homo sapien	360	39	81.2	1795	2	Q746C5_THET2	Q746c5	thermus the
288	39	81.2	373	2	Q7TXG5_MYCBO	Q7txg5	mycobacteri	361	39	81.2	1795	2	Q53VY7_THET8	Q53vy7	thermus the
289	39	81.2	373	2	O53253_MYCTU	O53253	mycobacteri	362	39	81.2	1878	2	Q66580_EEEV	Q66580	eastern equ
290	39	81.2	378	1	PBX4_MOUSE	Q99ne9	mus musculus	363	39	81.2	1951	2	Q6VZM1_CNPV	Q6vzm1	canarypox v
291	39	81.2	409	1	YNP2_YEAST	P53901	saccharomyc	364	39	81.2	1958	2	Q69340_9ALPH	Q69340	suid herpes
292	39	81.2	409	2	Q6Q5J6_YEAST	Q6q5j6	saccharomyc	365	39	81.2	2156	2	Q7SGT6_NEUCR	Q7sgt6	neurospora
293	39	81.2	423	2	Q4NSZ4_9DELT	Q4nsz4	anaeromyxob	366	39	81.2	2493	2	Q88789_EEEV	Q88789	eastern equ
294	39	81.2	424	2	Q6FP92_CANGA	Q6fp92	candida gla	367	39	81.2	2494	2	Q5XZF4_EEEV	Q5xxzf4	eastern equ
295	39	81.2	434	2	Q69TP3_ORYSA	Q69tp3	oryza sativ	368	39	81.2	2494	2	Q4QXJ8_EEEV	Q4qxj8	eastern equ
296	39	81.2	439	2	Q4NX07_9DELT	Q4nx07	anaeromyxob	369	39	81.2	2494	2	Q4QXK0_EEEV	Q4gxk0	eastern equ
297	39	81.2	440	1	BCL3_MOUSE	Q9z2f6	mus musculus	370	39	81.2	2511	2	Q95LC8_DIDMA	Q95lc8	didelphis m
298	39	81.2	444	2	Q9C197_AMAMU	Q9c197	amanita mus	371	39	81.2	2615	2	Q9CPI1_PASMU	Q9cpil	pasteurella
299	39	81.2	446	1	BCL3_HUMAN	P20749	homo sapien	372	39	81.2	2705	2	Q9W6V6_CHICK	Q9w6v6	gallus gall
300	39	81.2	459	2	Q8GMG2_STRGL	Q8gmg2	streptomyce	373	39	81.2	2725	2	Q9UKZ4_HUMAN	Q9ukz4	homo sapien
301	39	81.2	472	2	Q9RVZ5_DEIRA	Q9rvz5	deinococcus	374	39	81.2	2725	2	Q5JZ17_HUMAN	Q5jz17	homo sapien
302	39	81.2	477	2	Q6ZW60_HUMAN	Q6zw60	homo sapien	375	39	81.2	2731	2	Q9WTS4_MOUSE	Q9wts4	mus musculus
303	39	81.2	487	2	Q51KP3_MAGGR	Q51kp3	magnaporthe	376	39	81.2	2757	2	Q5CPF7_CRYPV	Q5cpf7	cryptospori
304	39	81.2	493	2	Q9LT85_ARATH	Q9lt85	arabidopsis	377	39	81.2	3027	2	Q6UDW5_PLAFA	Q6udw5	plasmodium
305	39	81.2	502	2	Q7R7A8_PLAYO	Q7r7a8	plasmodium	378	39	81.2	3596	1	RPOA_SHFV	Q68772	s replicase
306	39	81.2	506	2	Q504F6_BRARE	Q504f6	brachydanio	379	38	79.2	76	2	Q7S0D9_NEUCR	Q7s0d9	neurospora
307	39	81.2	519	2	Q91MK3_9PARA	Q91mk3	menangle vi	380	38	79.2	105	2	Q7EYT4_ORYSA	Q7eyt4	oryza sativ
308	39	81.2	523	2	Q93XJ2_PERFR	Q93xj2	perilla fru	381	38	79.2	115	2	Q72FB7_DESVH	Q72fb7	desulfovibr
309	39	81.2	532	2	Q6ZD77_ORYSA	Q6zd77	oryza sativ	382	38	79.2	117	2	Q67V80_ORYSA	Q67v80	oryza sativ
310	39	81.2	533	2	Q86ZC3_USTMA	Q86zc3	ustilago ma	383	38	79.2	121	1	AMEL_ORNAN	O97646	ornithorhyn
311	39	81.2	533	2	Q4PC32_USTMA	Q4pc32	ustilago ma	384	38	79.2	121	1	AMEL_TACAC	O97647	tachyglossu
312	39	81.2	536	2	Q8C8D2_MOUSE	Q8c8d2	mus musculus	385	38	79.2	135	2	Q4UUE9_XANCP	Q4uee9	xanthomonas
313	39	81.2	538	2	Q6C0T9_YARLI	Q6c0t9	yarrowia li	386	38	79.2	135	2	Q8P9E8_XANCP	Q8p9e8	xanthomonas
314	39	81.2	561	1	EFS_HUMAN	O43281	homo sapien	387	38	79.2	141	2	Q51E04_ENTHI	Q5be61	macaca fasc
315	39	81.2	580	2	Q7SEW0_NEUCR	Q7sew0	neurospora	388	38	79.2	154	2	Q5KNJ7_CRYNE	O5le04	entamoeba h
316	39	81.2	593	2	Q28248_CANFA	Q28248	canis famil	389	38	79.2	155	2	Q7XW07_ORYSA	Q7xw07	oryza sativ
317	39	81.2	600	2	Q866G0_CANFA	Q866g0	canis famil	390	38	79.2	156	2	Q82RS7_STRAW	Q82rs7	streptomyce
318	39	81.2	602	1	KCNA5_MOUSE	Q61762	mus musculus	391	38	79.2	178	2	Q7XHQ5_ORYSA	Q7xhg5	oryza sativ
319	39	81.2	602	2	Q9Z1R6_MOUSE	Q9z1r6	mus musculus	392	38	79.2	178	2	Q98680_9BETA	Q98680	simian cyto
320	39	81.2	609	2	Q4RFZ2_TETNG	Q4rfz2	tetraodon n	393	38	79.2	187	2	Q4LYX8_9BURK	Q4lyx8	burkholderi
321	39	81.2	612	2	Q503F8_BRARE	Q503f8	brachydanio	394	38	79.2	191	2	Q9A4I6_CAUCR	Q9a4i6	caulobacter
322	39	81.2	616	2	Q42200_BRARE	O42200	brachydanio	395	38	79.2	192	2	Q4I3J4_GIBZE	Q4i3j4	gibberella
323	39	81.2	636	2	Q7SF15_NEUCR	Q7sf15	neurospora	396	38	79.2					

397	38	79.2	192	2	Q53LW2_ORYSA	Q53lw2	oryza sativ	470	38	79.2	475	2	Q5RE17_PONPY	Q5rel17	pongo pygma
398	38	79.2	194	2	Q7PP14_ANOGA	Q7pp14	anopheles g	471	38	79.2	478	2	Q54LJ3_DICDI	Q54lj3	dictyosteli
399	38	79.2	200	2	Q8LAT8_ARATH	Q8lat8	arabidopsis	472	38	79.2	484	2	Q5YQH1_NOCFA	Q5yqh1	nocardia fa
400	38	79.2	200	2	Q9LT71_ARATH	Q9lt71	arabidopsis	473	38	79.2	485	1	T13H_TAXCU	Q8w4t9	taxus cuspi
401	38	79.2	204	2	Q72H82_THET2	Q72h82	thermus the	474	38	79.2	485	2	Q5BU48_9CONI	Q5bu48	taxus x med
402	38	79.2	204	2	Q5SMD8_THET8	Q5smd8	thermus the	475	38	79.2	485	2	Q56GD5_TAXCH	Q56gd5	taxus chine
403	38	79.2	207	2	Q98679_9BETA	Q98679	simian cyto	476	38	79.2	487	1	WASIP_RAT	Q6in36	rattus norv
404	38	79.2	208	2	Q6QI49_RAT	Q6qi49	rattus norv	477	38	79.2	489	2	Q4SFL7_TETNG	Q4sfl7	tetraodon n
405	38	79.2	216	2	Q6Z3X0_ORYSA	Q6z3x0	oryza sativ	478	38	79.2	491	2	Q67V14_ORYSA	Q67v14	oryza sativ
406	38	79.2	217	2	P93426_ORYSA	P93426	oryza sativ	479	38	79.2	495	2	Q6JD68_TAXCA	Q6jd68	taxus canad
407	38	79.2	217	2	Q5N952_ORYSA	Q5n952	oryza sativ	480	38	79.2	495	2	Q5S1U2_TAXCH	Q5s1u2	taxus chine
408	38	79.2	217	2	Q9M4Y5_ORYSA	Q9m4y5	oryza sativ	481	38	79.2	514	2	Q624I8_CAEBR	Q624i8	caenorhabdi
409	38	79.2	222	2	Q51B85_ENTHI	Q51b85	entamoeba h	482	38	79.2	527	1	TXK_HUMAN	P42681	homo sapien
410	38	79.2	222	2	Q9RXN1_DEIRA	Q9rxn1	deinococcus	483	38	79.2	527	1	TXK_MOUSE	P42682	mus musculu
411	38	79.2	229	2	Q6K9N3_ORYSA	Q6k9n3	oryza sativ	484	38	79.2	533	2	Q75IR3_ORYSA	Q75ir3	oryza sativ
412	38	79.2	230	2	O41979_MHV68	O41979	murid herpe	485	38	79.2	542	2	Q51UI7_MAGGR	Q51ui7	magnaporth
413	38	79.2	238	2	Q67VY6_ORYSA	Q67vy6	oryza sativ	486	38	79.2	544	2	Q6H082_FREDI	Q6h082	fremyella d
414	38	79.2	240	2	Q521N1_MAGGR	Q521n1	magnaporth	487	38	79.2	560	2	Q81YJ1_HUMAN	Q81yj1	homo sapien
415	38	79.2	247	2	Q6CSL7_YARLI	Q6csl7	yarrowia li	488	38	79.2	561	2	Q8G4W0_BIFLO	Q8g4w0	bifidobacte
416	38	79.2	256	2	Q6Z074_ORYSA	Q6z074	oryza sativ	489	38	79.2	570	2	Q5SS11_CRYNE	Q5ss11	cryptococcu
417	38	79.2	262	2	Q6ZRP0_HUMAN	Q6zrp0	homo sapien	490	38	79.2	585	2	Q5H067_XANOR	Q5h067	xanthomonas
418	38	79.2	262	2	Q6X263_9ALPH	Q6x263	bovine herp	491	38	79.2	585	2	Q7S6M3_NEUCR	Q7s6m3	neurospora
419	38	79.2	266	2	Q4IHX5_GIBZE	Q4ihx5	gibberella	492	38	79.2	590	2	Q6CDQ2_YARLI	Q6cdq2	yarrowia li
420	38	79.2	271	2	Q6D047_ERWCT	Q6d047	erwinia car	493	38	79.2	599	2	Q68FA5_XENTR	Q68fa5	xenopus tro
421	38	79.2	280	2	Q5SIL9_THET8	Q5eil9	thermus the	494	38	79.2	602	2	Q7SG00_NEUCR	Q7sg00	neurospora
422	38	79.2	280	2	Q72IZ6_THET2	Q72iz6	thermus the	495	38	79.2	605	1	PP2CD_HUMAN	O15297	homo sapien
423	38	79.2	282	2	Q7XGJ1_ORYSA	Q7xgj1	oryza sativ	496	38	79.2	605	2	Q53XP4_HUMAN	Q53xp4	homo sapien
424	38	79.2	282	2	Q94HM7_ORYSA	Q94hm7	oryza sativ	497	38	79.2	605	2	Q8NEA7_HUMAN	Q8nea7	homo sapien
425	38	79.2	287	2	Q7Y092_ORYSA	Q7y092	oryza sativ	498	38	79.2	605	2	Q6P991_HUMAN	Q6p991	homo sapien
426	38	79.2	292	2	Q4WSW5_ASPFU	Q4wsW5	aspergillus	499	38	79.2	605	2	Q7SYD2_BRARE	Q7syd2	brachydanio
427	38	79.2	292	2	Q6MYZ3_ASPFU	Q6myz3	aspergillus	500	38	79.2	612	2	Q4SN07_TETNG	Q4sn07	tetraodon n
428	38	79.2	300	2	Q4NPH7_9DELT	Q4nph7	anaeromyxob	501	38	79.2	624	2	Q87A69_XYLFT	Q87a69	xylella fas
429	38	79.2	303	2	Q7YV35_9TRYP	Q7yv35	trypanosoma	502	38	79.2	636	2	Q60I21_EPTBU	Q60i21	eptatretus
430	38	79.2	304	2	Q876G4_SACBA	Q876g4	saccharomyc	503	38	79.2	639	1	ZIC5_HUMAN	Q96t25	homo sapien
431	38	79.2	305	2	Q6YWH3_ORYSA	Q6ywh3	oryza sativ	504	38	79.2	639	2	Q5VYB0_HUMAN	Q5vyb0	homo sapien
432	38	79.2	306	2	Q4RUA7_TETNG	Q4rua7	tetraodon n	505	38	79.2	644	2	Q7RIV9_PLAYO	Q7riv9	plasmodium
433	38	79.2	308	1	MPASA_LOLPR	Q40240	lolium pere	506	38	79.2	649	2	Q9UIM6_DROME	Q9ulm6	drosophila
434	38	79.2	318	1	ADT2_YEAST	P18239	saccharomyc	507	38	79.2	653	2	Q6UPR4_CHLRE	Q6upr4	chlamydomon
435	38	79.2	324	2	Q4STL5_TETNG	Q4stl5	tetraodon n	508	38	79.2	656	2	Q7S6U6_NEUCR	Q7s6u6	neurospora
436	38	79.2	330	2	Q5QNI5_ORYSA	Q5qni5	oryza sativ	509	38	79.2	657	1	BTk_CHICK	Q8jh64	gallus gall
437	38	79.2	331	2	Q8NRR3_CORGL	Q8nrr3	corynebacte	510	38	79.2	657	2	Q5ZLF2_CHICK	Q5zlf2	gallus gall
438	38	79.2	336	2	Q848F5_STRLI	Q848f5	streptomyce	511	38	79.2	665	2	Q5KGX2_CRYNE	Q5kgx2	cryptococcu
439	38	79.2	339	2	Q524X6_MAGGR	Q524x6	magnaporth	512	38	79.2	665	2	Q50XG4_ENTHI	Q50xg4	entamoeba h
440	38	79.2	344	2	Q5T8S7_HUMAN	Q5t8s7	homo sapien	513	38	79.2	668	1	PBS2_YEAST	P08018	saccharomyc
441	38	79.2	349	2	Q87X32_PBSM	Q87x32	pseudomonas	514	38	79.2	677	2	Q6FL74_CANGA	Q6fl74	candida gla
442	38	79.2	356	2	Q924W1_RAT	Q924w1	rattus norv	515	38	79.2	681	2	Q6BUS0_DEBHA	Q6bus0	debaryomyce
443	38	79.2	360	2	Q719M0_BRAPL	Q719m0	branchiosto	516	38	79.2	683	2	Q96W50_DEBHA	Q96w50	debaryomyce
444	38	79.2	364	2	Q55QI6_CRYNE	Q55qi6	cryptococcu	517	38	79.2	687	2	Q7QDV6_ANOGA	Q7qdv6	anopheles g
445	38	79.2	364	2	Q5KFNO_CRYNE	Q5kfn0	cryptococcu	518	38	79.2	687	2	Q51G68_ENTHI	Q51g68	entamoeba h
446	38	79.2	367	2	Q746C6_THET2	Q746c6	thermus the	519	38	79.2	691	2	Q755N1_ASHGO	Q755n1	ashbya goss
447	38	79.2	374	2	Q7VXS4_BORPE	Q7vxs4	bordetella	520	38	79.2	705	2	Q7RYU0_NEUCR	Q7ryu0	neurospora
448	38	79.2	374	2	Q7W2T8_BORPA	Q7w2t8	bordetella	521	38	79.2	707	2	Q8CHU0_MOUSE	Q8chu0	mus musculu
449	38	79.2	374	2	Q7WDT6_BORBR	Q7wdt6	bordetella	522	38	79.2	708	2	Q8SX98_DROME	Q8sx98	drosophila
450	38	79.2	387	2	Q84HB1_STRCZ	Q84hb1	streptomyce	523	38	79.2	724	2	Q6CN49_KLULA	Q6cn49	kluyveromyc
451	38	79.2	398	2	Q8VBA1_WSSV	Q8vba1	white spot	524	38	79.2	738	2	Q9Z213_RAT	Q9z213	rattus norv
452	38	79.2	398	2	Q91LI1_WSSV	Q91li1	white spot	525	38	79.2	740	2	Q5TTF1_ANOGA	Q5ttf1	anopheles g
453	38	79.2	400	2	Q8FQS2_COREF	Q8fqs2	corynebacte	526	38	79.2	751	2	Q8MLU0_DROME	Q8mlu0	drosophila
454	38	79.2	402	2	Q4RX44_TETNG	Q4rx44	tetraodon n	527	38	79.2	753	2	Q4PA88_USTMA	Q4pa88	ustilago ma
455	38	79.2	403	2	Q4SPM0_TETNG	Q4spm0	tetraodon n	528	38	79.2	761	2	Q7KVL6_DROME	Q7kv16	drosophila
456	38	79.2	404	2	Q8X095_NEUCR	Q8x095	neurospora	529	38	79.2	774	2	Q4RRS9_TETNG	Q4rre9	tetraodon n
457	38	79.2	407	2	Q744V4_MYCPA	Q744v4	mycobacteri	530	38	79.2	787	2	Q63X79_BURPS	Q63x79	burkholderi
458	38	79.2	407	2	Q9AA59_CAUCR	Q9aa59	caulobacter	531	38	79.2	787	2	Q62MM5_BURMA	Q62mm5	burkholderi
459	38	79.2	408	1	YHFW_ECOLI	P45549	escherichia	532	38	79.2	837	2	Q60Z34_CAEBR	Q60z34	caenorhabdi
460	38	79.2	408	2	Q8FCW4_ECOL6	Q8fcw4	escherichia	533	38	79.2	870	2	Q7VZJ7_BORPE	Q7vzj7	bordetella
461	38	79.2	408	2	Q83PX2_SHIFL	Q83px2	shigella fl	534	38	79.2	870	2	Q7WHV9_BORBR	Q7whv9	bordetella
462	38	79.2	418	2	Q9KYZ8_STRCO	Q9kyz8	streptomyce	535	38	79.2	874	2	Q5JU58_HUMAN	Q5ju58	homo sapien
463	38	79.2	425	2	Q9C0Q7_NEUCR	Q9c0q7	neurospora	536	38	79.2	876	2	Q6XLI5_BRARE	Q6xli5	brachydanio
464	38	79.2	427	2	Q5ASC8_EMENI	Q5asc8	aspergillus	537	38	79.2	883	2	Q4RUA9_TETNG	Q4rua9	tetraodon n
465	38	79.2	430	2	Q8IVR6_HUMAN	Q8ivr6	homo sapien	538	38	79.2	884	2	Q5NBA4_ORYSA	Q5nba4	oryza sativ
466	38	79.2	439	2	Q619S9_CAEBR	Q619s9	caenorhabdi	539	38	79.2	885	2	Q4PHH1_USTMA	Q4phh1	ustilago ma
467	38	79.2	443	1	SAM68_MOUSE	Q60749	m kh domain	540	38	79.2	888	2	Q5B371_EMENI	Q5b371	aspergillus
468	38	79.2	443	1	SAM68_MOUSE	Q91v33	r kh domain	541	38	79.2	894	2	Q7W9W8_BORPA	Q7w9w8	bordetella
469	38	79.2	474	2	Q5B433_EMENI	Q5b433	aspergillus	542	38	79.2	901	2	Q53VY8_THET8	Q53vy8	thermus the

543	38	79.2	922	2	Q8UVF2_COTCO	Q8uvf2	coturnix co
544	38	79.2	925	1	PIP1_YEAST	P40020	saccharomyc
545	38	79.2	925	2	Q4IRH1_GIBZE	Q4irh1	gibberella
546	38	79.2	949	2	Q5JUX1_HUMAN	Q5jux1	homo sapien
547	38	79.2	995	1	MA2B2_PIG	Q28949	sus scrofa
548	38	79.2	1003	2	Q9C235_NEUCR	Q9c235	neurospora
549	38	79.2	1006	2	Q55R68_CRYNE	Q55r68	cryptococcu
550	38	79.2	1006	2	Q5KF02_CRYNE	Q5kf02	cryptococcu
551	38	79.2	1009	2	Q55R69_CRYNE	Q55r69	cryptococcu
552	38	79.2	1009	2	Q5KF01_CRYNE	Q5kf01	cryptococcu
553	38	79.2	1009	2	Q4QEU8_LEIMA	Q4qeu8	leishmania
554	38	79.2	1057	2	Q6CES9_YARLI	Q6ces9	yarrowia li
555	38	79.2	1101	2	Q5K9C7_CRYNE	Q5k9c7	cryptococcu
556	38	79.2	1101	2	Q55K74_CRYNE	Q55k74	cryptococcu
557	38	79.2	1125	2	Q5BFB7_EMENI	Q5bfb7	aspergillus
558	38	79.2	1134	2	Q80TS1_MOUSE	Q80ts1	mus musculus
559	38	79.2	1157	1	DXH37_HUMAN	Q8iy37	homo sapien
560	38	79.2	1169	2	Q5KIS4_CRYNE	Q5kis4	cryptococcu
561	38	79.2	1169	2	Q55TL1_CRYNE	Q55tl1	cryptococcu
562	38	79.2	1218	2	Q9V6T7_DROME	Q9v6t7	drosophila
563	38	79.2	1229	2	Q4WS24_ASPFU	Q4ws24	aspergillus
564	38	79.2	1236	2	Q6MY63_ASPFU	Q6my63	aspergillus
565	38	79.2	1379	2	Q8I7P4_DICDI	Q8i7p4	dictyosteli
566	38	79.2	1379	2	Q54D97_DICDI	Q54d97	dictyosteli
567	38	79.2	1393	2	Q55U57_CRYNE	Q55u57	cryptococcu
568	38	79.2	1398	2	Q7F1V5_ORYSA	Q7flv5	oryza sativ
569	38	79.2	1402	2	Q5KI87_CRYNE	Q5ki87	cryptococcu
570	38	79.2	1478	2	Q5JU85_HUMAN	Q5ju85	homo sapien
571	38	79.2	1560	2	Q60275_HUMAN	Q60275	homo sapien
572	38	79.2	1677	2	Q4RVQ9_TETNG	Q4rvq9	tetraodon n
573	38	79.2	1700	2	Q9VJU9_DROME	Q9vjj9	drosophila
574	38	79.2	1914	2	Q8X0I0_NEUCR	Q8x0i0	neurospora
575	38	79.2	1919	2	Q7RW47_NEUCR	Q7rw47	neurospora
576	38	79.2	2025	2	Q4X1U1_ASPFU	Q4x1u1	aspergillus
577	38	79.2	2039	2	Q7S3G9_NEUCR	Q7s3g9	neurospora
578	38	79.2	2161	1	SHAN1_HUMAN	Q9y566	homo sapien
579	38	79.2	2295	2	Q7S2P2_NEUCR	Q7s2p2	neurospora
580	38	79.2	2837	2	Q96U29_NEUCR	Q96u29	neurospora
581	38	79.2	3980	2	Q4T089_TETNG	Q4t089	tetraodon n
582	37.5	78.1	71	2	Q6MRN1_BDEBA	Q6mrn1	bdellovibri
583	37	77.1	52	2	Q6SPR5_SHEEP	Q6spr5	ovis aries
584	37	77.1	55	2	Q9TJU6_CERNI	Q9tju6	cervus nipp
585	37	77.1	60	2	Q6SPR6_CEREL	Q6spr6	cervus elap
586	37	77.1	75	2	Q4SHC6_TETNG	Q4shc6	tetraodon n
587	37	77.1	79	2	Q9PX71_9BETA	Q9px71	human herpe
588	37	77.1	83	2	Q4GYV7_9TRYP	Q4gyv7	trypanosoma
589	37	77.1	93	2	Q58N36_9CAUD	Q58n36	cytaphage
590	37	77.1	99	2	Q5SMJ8_ORYSA	Q5smj8	oryza sativ
591	37	77.1	104	2	Q9YBH3_AERPE	Q9ybh3	aeropyrum p
592	37	77.1	104	2	Q71RB1_HUMAN	Q71rb1	homo sapien
593	37	77.1	106	2	Q693N3_CLOAL	Q693n3	clogmia alb
594	37	77.1	107	2	Q7EZ89_ORYSA	Q7ez89	oryza sativ
595	37	77.1	109	2	Q9SY21_ARATH	Q9sy21	arabidopsis
596	37	77.1	123	2	Q8BNN7_MOUSE	Q8bnn7	mus musculu
597	37	77.1	124	2	Q92UK1_RHIME	Q92uk1	rhizobium m
598	37	77.1	126	2	Q67VU5_ORYSA	Q67vu5	oryza sativ
599	37	77.1	129	2	Q72IL0_THET2	Q72il0	thermus the
600	37	77.1	143	2	Q7PNB6_ANOGA	Q7pnb6	anopheles g
601	37	77.1	151	2	Q84PY1_ORYSA	Q84py1	oryza sativ
602	37	77.1	151	2	Q4NAA2_9MICC	Q4naa2	arthrobacte
603	37	77.1	162	2	Q6ZP66_HUMAN	Q6zp66	homo sapien
604	37	77.1	166	2	Q9X8W3_STRCO	Q9x8w3	streptomyce
605	37	77.1	167	2	Q4PMF1_IXOSC	Q4pmf1	ixodes scap
606	37	77.1	175	2	Q83739_9LUTE	Q83739	beet mild y
607	37	77.1	181	2	Q5P0T0_AZOSE	Q5p0t0	azocarcus sp
608	37	77.1	189	2	Q6Z6J1_ORYSA	Q6z6j1	oryza sativ
609	37	77.1	194	2	Q7NCA5_GLOVI	Q7nca5	gloeobacter
610	37	77.1	196	2	Q69TM5_ORYSA	Q69tm5	oryza sativ
611	37	77.1	202	2	Q75KQ8_ORYSA	Q75kq8	oryza sativ
612	37	77.1	203	2	Q9RCX9_STRCO	Q9rcx9	streptomyce
613	37	77.1	206	2	Q7NUT6_CHRVO	Q7nut6	chromobacte
614	37	77.1	210	2	Q9VEH6_DROME	Q9veh6	drosophila
615	37	77.1	211	2	Q55UJ5_CRYNE	Q55uj5	cryptococcu

Q5khv0	cryptococcu	211	2	Q5KHV0_CRYNE	Q5khv0	cryptococcu
Q5lh25	entamoeba h	211	2	Q51H25_ENTHI	Q5lh25	entamoeba h
Q60ex0	oryza sativ	220	2	Q60EX0_ORYSA	Q60ex0	oryza sativ
Q6z9l5	oryza sativ	220	2	Q6Z9L5_ORYSA	Q6z9l5	oryza sativ
Q55tl7	cryptococcu	223	2	Q55TL7_CRYNE	Q55tl7	cryptococcu
Q5kir8	cryptococcu	223	2	Q5KIR8_CRYNE	Q5kir8	cryptococcu
Q60ee3	oryza sativ	228	2	Q60EE3_ORYSA	Q60ee3	oryza sativ
Q6yzn8	oryza sativ	232	2	Q6YZN8_ORYSA	Q6yzn8	oryza sativ
Q6aw70	bombyx mori	237	2	Q6AW70_9VIRU	Q6aw70	bombyx mori
Q5ysd4	nocardia fa	244	2	Q5YSD4_NOCFA	Q5ysd4	nocardia fa
Q63lf2	burkholderi	245	2	Q63LF2_BURPS	Q63lf2	burkholderi
Q7zyi9	xenopus lae	250	2	Q7ZYI9_XENLA	Q7zyi9	xenopus lae
Q6u536	hepatitis c	254	2	Q6U536_9HEPC	Q6u536	hepatitis c
Q6u541	hepatitis c	254	2	Q6U541_9HEPC	Q6u541	hepatitis c
Q7lbb7	brucella ab	256	2	Q71BB7_BRUAB	Q7lbb7	brucella ab
Q32585	escherichia	258	2	Q32585_ECOLI	Q32585	escherichia
Q8gj13	escherichia	258	2	Q8GJ13_ECOLI	Q8gj13	escherichia
Q5gx23	xanthomonas	263	2	Q5GX23_XANOR	Q5gx23	xanthomonas
Q5gx32	xanthomonas	263	2	Q5GX32_XANOR	Q5gx32	xanthomonas
Q524i5	magnaporthe	265	2	Q524I5_MAGGR	Q524i5	magnaporthe
Q9bhf8	leishmania	267	2	Q9BHF8_LEIMA	Q9bhf8	leishmania
Q67w47	oryza sativ	271	2	Q67W47_ORYSA	Q67w47	oryza sativ
Q8bjs5	mus musculu	272	2	Q8BJS5_MOUSE	Q8bjs5	mus musculu
Q82hn1	streptomyce	273	2	Q82HN1_STRAW	Q82hn1	streptomyce
Q72u85	leptospiro	276	2	Q72U85_LEPIC	Q72u85	leptospiro
Q8f0v1	leptospiro	276	2	Q8F0V1_LEPIN	Q8f0v1	leptospiro
Q4tpv6	erythrobact	279	2	Q4TPV6_9SPHN	Q4tpv6	erythrobact
Q5ql13	oryza sativ	280	2	Q5QL13_ORYSA	Q5ql13	oryza sativ
Q6z6e2	oryza sativ	281	2	Q6Z6E2_ORYSA	Q6z6e2	oryza sativ
P78977	yarrowia li	285	2	P78977_YARLI	P78977	yarrowia li
Q67tr5	oryza sativ	286	2	Q67TR5_ORYSA	Q67tr5	oryza sativ
Q4wuc2	aspergillus	287	2	Q4WUC2_ASPFU	Q4wuc2	aspergillus
Q67we9	oryza sativ	287	2	Q67WE9_ORYSA	Q67we9	oryza sativ
Q5dts1	mus musculu	287	2	Q5DTS1_MOUSE	Q5dts1	mus musculu
Q5z572	oryza sativ	289	2	Q5Z572_ORYSA	Q5z572	oryza sativ
Q7eyb7	oryza sativ	289	2	Q7EYB7_ORYSA	Q7eyb7	oryza sativ
Q9sxe7	arabidopsis	297	2	Q9SXE7_ARATH	Q9sxe7	arabidopsis
Q9rvv7	deinococcus	298	2	Q9RVV7_DEIRA	Q9rvv7	deinococcus
Q93f63	bacillus sp	299	2	Q93F63_9BACI	Q93f63	bacillus sp
Q8lrl5	petunia hyb	304	2	Q8LRL5_PETHY	Q8lrl5	petunia hyb
Q4qy52	simian t-ly	304	2	Q4QY52_9STL1	Q4qy52	simian t-ly
Q787c4	neurospora	313	2	Q7S7C4_NEUCR	Q787c4	neurospora
Q548l7	mus musculu	314	2	Q548L7_MOUSE	Q548l7	mus musculu
Q6mw23	neurospora	317	2	Q6MW23_NEUCR	Q6mw23	neurospora
Q4zlh4	pseudomonas	318	2	Q4ZLH4_PSESY	Q4zlh4	pseudomonas
Q53sb7	homo sapien	321	2	Q53SB7_HUMAN	Q53sb7	homo sapien
Q70k98	gordonia we	322	2	Q70K98_9ACTO	Q70k98	gordonia we
Q9lk80	spodoptera	322	2	Q91K80_NPVST	Q9lk80	spodoptera
Q6mes3	parachlamyd	323	2	Q6MES3_PARUW	Q6mes3	parachlamyd
Q8lni0	oryza sativ	343	2	Q8LNI0_ORYSA	Q8lni0	oryza sativ
Q4zlh3	pseudomonas	343	2	Q4ZLH3_PSESY	Q4zlh3	pseudomonas
Q8d8u2	vibrio vuln	343	2	Q8D8U2_VIBVU	Q8d8u2	vibrio vuln
Q7ngk0	gloeobacter	346	2	Q7NGK0_GLOVI	Q7ngk0	gloeobacter
Q7mln2	vibrio vuln	346	2	Q7MLN2_VIBVY	Q7mln2	vibrio vuln
Q4s820	tetraodon n	348	2	Q4S820_TETNG	Q4s820	tetraodon n
Q4zpb39	pseudomonas	349	2	Q4ZPB39_PSESY	Q4zpb39	pseudomonas
Q55l60	cryptococcu	355	2	Q55L60_CRYNE	Q55l60	cryptococcu
Q5kac1	cryptococcu	355	2	Q5KAC1_CRYNE	Q5kac1	cryptococcu
Q9vji6	drosophila	357	2	Q9VJI6_DROME	Q9vji6	drosophila
Q4tfe3	tetraodon n	362	2	Q4TFE3_TETNG	Q4tfe3	tetraodon n
Q7xg36	oryza sativ	368	2	Q7XG36_ORYSA	Q7xg36	oryza sativ
Q94i20	oryza sativ	368	2	Q94I20_ORYSA	Q94i20	oryza sativ
Q827m8	streptomyce	370	2	Q827M8_STRAW	Q827m8	streptomyce
Q529g0	magnaporthe	372	2	Q529G0_MAGGR	Q529g0	magnaporthe
Q9nps0	homo sapien	375	2	Q9NPS0_HUMAN	Q9nps0	homo sapien
Q9h7e6	homo sapien	376	2	Q9H7E6_HUMAN	Q9h7e6	homo sapien
Q4ql77	leishmania	376	2	Q4QL77_LEIMA	Q4ql77	leishmania
Q88933	mus musculu	382	2	Q88933_MOUSE	Q88933	mus musculu
Q7yle1	oryza sativ	383	2	Q7Y1E1_ORYSA	Q7yle1	oryza sativ
Q9fpq5	chlamydomon	386	2	Q9FPQ5_CHLRE	Q9fpq5	chlamydomon
Q9z2u3	mus musculu	387	2	Q9Z2U3_MOUSE	Q9z2u3	mus musculu
Q4r9f1	macaca fasc	392	2	Q4R9F1_MACFA	Q4r9f1	macaca fasc
Q8k437	mus musculu	397	2	Q8K437_MOUSE	Q8k437	mus musculu

689	37	77.1	397	2	Q8K0J2_MOUSE	Q8k0j2 mus musculu
690	37	77.1	397	2	Q6PA02_MOUSE	Q6pa02 mus musculu
691	37	77.1	400	2	Q73VH2_MYCPA	Q73vh2 mycobacteri
692	37	77.1	408	2	O44776_CABEL	O44776 caenorhabdi
693	37	77.1	409	2	Q9SBM1_VOLCA	Q9sbm1 volvox cart
694	37	77.1	411	2	Q7SE19_NEUCR	Q7se19 neurospora
695	37	77.1	413	2	O15417_HUMAN	O15417 homo sapien
696	37	77.1	414	2	Q55LS8_CRYNE	Q55ls8 cryptococcu
697	37	77.1	414	2	Q5K8Z7_CRYNE	Q5k8z7 cryptococcu
698	37	77.1	414	2	Q8NBX7_HUMAN	Q8nbx7 homo sapien
699	37	77.1	417	2	Q6PIN9_HUMAN	Q6pin9 homo sapien
700	37	77.1	426	2	Q5YUA8_NOCFA	Q5yua8 nocardia fa
701	37	77.1	429	2	Q5BFA3_EMENI	Q5bfa3 aspergillus
702	37	77.1	430	2	Q5TQU9_ANOGA	Q5tqu9 anopheles g
703	37	77.1	435	1	RAD52_ASHGO	Q756f4 ashbya goss
704	37	77.1	445	2	Q95KJ7_MACFA	Q95kj7 macaca fasc
705	37	77.1	446	2	Q5TTT0_ANOGA	Q5ttt0 anopheles g
706	37	77.1	454	2	Q9RVH3_DEIRA	Q9rvh3 deinococcus
707	37	77.1	471	2	Q9NPM5_HUMAN	Q9npm5 homo sapien
708	37	77.1	481	2	Q82CH7_STRAW	Q82ch7 streptomyce
709	37	77.1	485	2	Q7UA85_SYNPX	Q7ua85 synechococc
710	37	77.1	487	1	EBN2_EBV	P12978 epstein-bar
711	37	77.1	487	2	Q777H1_9GAMA	Q777h1 human herpe
712	37	77.1	488	2	Q6CBY3_YARLI	Q6cby3 yarrowia li
713	37	77.1	488	2	Q95UR9_SCHAM	Q95ur9 schistocerc
714	37	77.1	488	2	Q4Q2C0_LEIMA	Q4q2c0 leishmania
715	37	77.1	489	2	Q873K2_NEUCR	Q873k2 neurospora
716	37	77.1	489	2	Q73XK9_MYCPA	Q73xk9 mycobacteri
717	37	77.1	490	2	Q69023_9GAMA	Q69023 human herpe
718	37	77.1	499	2	Q6YTE2_ORYSA	Q6yte2 oryza sativ
719	37	77.1	500	2	Q69X58_ORYSA	Q69x58 oryza sativ
720	37	77.1	500	2	Q6YTF1_ORYSA	Q6ytf1 oryza sativ
721	37	77.1	514	2	Q8X0Q3_NEUCR	Q8x0q3 neurospora
722	37	77.1	519	2	Q4I9W3_GIBZE	Q4i9w3 gibberella
723	37	77.1	520	2	Q8R4X9_MOUSE	Q8r4x9 mus musculu
724	37	77.1	520	2	Q99MY6_MOUSE	Q99my6 mus musculu
725	37	77.1	521	2	Q73XJ0_MYCPA	Q73xj0 mycobacteri
726	37	77.1	521	2	Q99P73_MOUSE	Q99p73 mus musculu
727	37	77.1	521	2	Q8VDL9_MOUSE	Q8vdl9 mus musculu
728	37	77.1	524	2	Q9BZE0_HUMAN	Q9bze0 homo sapien
729	37	77.1	535	2	Q55Y03_CRYNE	Q55y03 cryptococcu
730	37	77.1	535	2	Q5KLY7_CRYNE	Q5kly7 cryptococcu
731	37	77.1	542	2	Q55J32_CRYNE	Q55j32 cryptococcu
732	37	77.1	547	2	Q9CA82_ARATH	Q9ca82 arabidopsis
733	37	77.1	551	2	Q9NP74_HUMAN	Q9np74 homo sapien
734	37	77.1	555	1	GPI_CHLRE	Q9fpq6 chlamydomon
735	37	77.1	563	2	Q613F0_CABBR	Q613f0 caenorhabdi
736	37	77.1	564	2	Q7SCU8_NEUCR	Q7scu8 neurospora
737	37	77.1	569	2	Q9CU24_MOUSE	Q9cu24 mus musculu
738	37	77.1	580	2	Q51UJ2_MAGGR	Q51uj2 magnaporthe
739	37	77.1	582	1	MNT_HUMAN	Q99583 homo sapien
740	37	77.1	582	2	Q6MVT9_NEUCR	Q6mvt9 neurospora
741	37	77.1	582	2	Q5N5V7_SYNPF6	Q5n5v7 synechococc
742	37	77.1	589	2	Q8RW91_ARATH	Q8rw91 arabidopsis
743	37	77.1	591	1	MNT_MOUSE	O08789 mus musculu
744	37	77.1	591	2	Q6GTJ3_MOUSE	Q6gtj3 mus musculu
745	37	77.1	593	2	Q9MA04_ARATH	Q9ma04 arabidopsis
746	37	77.1	599	2	Q80VL9_MOUSE	Q80vl9 mus musculu
747	37	77.1	604	2	Q9L252_STRCO	Q9l252 streptomyce
748	37	77.1	608	2	Q4H2Q7_CIOIN	Q4h2q7 ciona intes
749	37	77.1	610	2	Q6ZC88_ORYSA	Q6zc88 oryza sativ
750	37	77.1	612	2	Q8FSI9_COREF	Q8fsi9 corynebacte
751	37	77.1	616	2	Q8T0J3_DROME	Q8t0j3 drosophila
752	37	77.1	633	2	Q5Z5T3_ORYSA	Q5z5t3 oryza sativ
753	37	77.1	640	2	Q60W81_CABBR	Q60w81 caenorhabdi
754	37	77.1	645	2	Q7RWI8_NEUCR	Q7rwi8 neurospora
755	37	77.1	646	1	FBX46_HUMAN	Q6pj61 homo sapien
756	37	77.1	647	2	Q96FT7_HUMAN	Q96ft7 homo sapien
757	37	77.1	648	2	Q5GYL8_XANOR	Q5gy18 xanthomonas
758	37	77.1	648	2	Q4UTQ5_XANCP	Q4utq5 xanthomonas
759	37	77.1	648	2	Q8P9Y0_XANCP	Q8p9y0 xanthomonas
760	37	77.1	648	2	Q8PLQ7_XANAC	Q8plq7 xanthomonas
761	37	77.1	657	2	Q7N7J1_PHOLL	Q7n7j1 photorhabdu

762	37	77.1	663	1	TERM_ADEB3	O55439 bovine aden
763	37	77.1	665	2	Q8XR57_RALSO	Q8xr57 ralstonia s
764	37	77.1	669	2	Q8HJ8_ORYSA	Q8hj8 oryza sativ
765	37	77.1	676	1	EXTL1_HUMAN	Q92935 homo sapien
766	37	77.1	676	2	Q6GSC1_HUMAN	Q6gsc1 homo sapien
767	37	77.1	686	2	Q5B4D7_EMENI	Q5b4d7 aspergillus
768	37	77.1	688	2	Q871U3_NEUCR	Q871u3 neurospora
769	37	77.1	691	2	Q55S15_CRYNE	Q55s15 cryptococcu
770	37	77.1	691	2	Q5KGX4_CRYNE	Q5kgx4 cryptococcu
771	37	77.1	693	2	Q4P3N9_USTMA	Q4p3n9 ustilago ma
772	37	77.1	695	2	Q9XUQ0_CABEL	Q9xuq0 caenorhabdi
773	37	77.1	697	2	Q61K69_CABBR	Q61k69 caenorhabdi
774	37	77.1	700	2	Q8N0R5_BOMMO	Q8n0r5 bombyx mori
775	37	77.1	703	2	Q9BYI8_HUMAN	Q9byi8 homo sapien
776	37	77.1	703	2	Q8N6Y0_HUMAN	Q8n6y0 homo sapien
777	37	77.1	724	2	Q8BWD1_MOUSE	Q8bwd1 mus musculu
778	37	77.1	745	2	Q9LFT3_ARATH	Q9lft3 arabidopsis
779	37	77.1	752	2	Q90ZP4_ONCMY	Q90zp4 oncorhynchu
780	37	77.1	757	2	Q4NRX0_9DELT	Q4nrx0 anaeromymbob
781	37	77.1	760	2	Q66BK5_YERPS	Q66bk5 yersinia ps
782	37	77.1	767	2	Q4NZM3_9DELT	Q4nzm3 anaeromymbob
783	37	77.1	776	2	Q6XK32_ORYSA	Q6xk32 oryza sativ
784	37	77.1	777	2	Q7S7K1_NEUCR	Q7s7k1 neurospora
785	37	77.1	780	2	Q4N6E3_THEPA	Q4n6e3 theileria p
786	37	77.1	782	2	Q4WEV6_ASPFU	Q4wev6 aspergillus
787	37	77.1	784	2	Q53SQ1_HUMAN	Q53sq1 homo sapien
788	37	77.1	784	2	Q86NX9_DROME	Q86nx9 drosophila
789	37	77.1	786	2	Q5TTK2_ANOGA	Q5ttk2 anopheles g
790	37	77.1	794	2	Q7RYS4_NEUCR	Q7rys4 neurospora
791	37	77.1	821	2	Q5KCM4_CRYNE	Q5kcm4 cryptococcu
792	37	77.1	825	2	Q6GN93_XENLA	Q6gn93 xenopus lae
793	37	77.1	830	2	Q95UR8_SCHAM	Q95ur8 schistocerc
794	37	77.1	836	2	Q4RYH6_TETNG	Q4ryh6 tetraodon n
795	37	77.1	838	2	Q8R3S1_MOUSE	Q8r3s1 mus musculu
796	37	77.1	845	2	Q7S4M4_NEUCR	Q7s4m4 neurospora
797	37	77.1	851	2	Q66BA9_YERPS	Q66ba9 yersinia ps
798	37	77.1	863	2	Q9LEE8_MAIZE	Q9lee8 zea mays (m
799	37	77.1	867	2	Q4X0Y7_ASPFU	Q4x0y7 aspergillus
800	37	77.1	871	2	Q6DJB9_XENTR	Q6djb9 xenopus tro
801	37	77.1	876	2	Q9P4Y9_NEUCR	Q9p4y9 neurospora
802	37	77.1	880	1	PHTF_DROME	Q9v9a8 drosophila
803	37	77.1	880	2	Q614D5_CABBR	Q614d5 caenorhabdi
804	37	77.1	885	2	Q55KS5_CRYNE	Q55ks5 cryptococcu
805	37	77.1	885	2	Q5KAQ0_CRYNE	Q5kaq0 cryptococcu
806	37	77.1	908	2	Q80VR4_MOUSE	Q80vyr4 mus musculu
807	37	77.1	916	2	Q59LR9_CANAL	Q59lr9 candida alb
808	37	77.1	936	2	Q8QRV7_9BETA	Q8qrv7 pongine her
809	37	77.1	939	2	Q4SPD1_TETNG	Q4spd1 tetraodon n
810	37	77.1	943	2	Q7SDY9_NEUCR	Q7sdy9 neurospora
811	37	77.1	951	2	Q7RUP0_NEUCR	Q7rup0 neurospora
812	37	77.1	970	2	Q60L67_CABBR	Q60l67 caenorhabdi
813	37	77.1	971	1	R3HDM_HUMAN	Q15032 homo sapien
814	37	77.1	1005	2	Q8MQ54_CABEL	Q8mq54 caenorhabdi
815	37	77.1	1011	2	Q9P944_PNECA	Q9p944 pneumocysti
816	37	77.1	1014	2	Q6DRC5_BRARE	Q6drc5 brachydanio
817	37	77.1	1019	2	Q4RCG15_TETNG	Q4rg15 tetraodon n
818	37	77.1	1024	2	Q4WXM8_ASPFU	Q4wxm8 aspergillus
819	37	77.1	1024	2	Q9XGS7_CHLRE	Q9xgs7 chlamydomon
820	37	77.1	1061	2	Q5R7K5_PONPY	Q5r7k5 pongo pygma
821	37	77.1	1062	2	Q4Q9P4_LEIMA	Q4q9p4 leishmania
822	37	77.1	1063	2	Q4VE09_RUBV	Q4ve09 rubella vir
823	37	77.1	1069	2	Q52BT2_MAGGR	Q52bt2 magnaporthe
824	37	77.1	1099	2	Q8IW32_HUMAN	Q8iw32 homo sapien
825	37	77.1	1116	2	Q7PLF4_DROME	Q7plf4 drosophila
826	37	77.1	1126	2	Q641I5_XENLA	Q641i5 xenopus lae
827	37	77.1	1140	2	Q20Z99_CABEL	Q20z99 caenorhabdi
828	37	77.1	1146	2	Q6PLP6_CHLRE	Q6plp6 chlamydomon
829	37	77.1	1147	1	SFR15_HUMAN	Q95104 homo sapien
830	37	77.1	1147	2	Q5D4Z9_9EURO	Q5d4z9 monascus pi
831	37	77.1	1147	2	Q5D500_9EURO	Q5d500 monascus pi
832	37	77.1	1154	2	Q5H2U5_XANOR	Q5h2u5 xanthomonas
833	37	77.1	1154	2	Q8PIN7_XANAC	Q8pin7 xanthomonas
834	37	77.1	1155	2	Q4UWT9_XANCP	Q4uwt9 xanthomonas

689	37	77.1	397	2	Q8K0J2_MOUSE	Q8k0j2 mus musculu
690	37	77.1	397	2	Q6PA02_MOUSE	Q6pa02 mus musculu
691	37	77.1	400	2	Q73VH2_MYCPA	Q73vh2 mycobacteri
692	37	77.1	408	2	O44776_CABEL	O44776 caenorhabdi
693	37	77.1	409	2	Q9SBM1_VOLCA	Q9sbm1 volvox cart
694	37	77.1	411	2	Q7SE19_NEUCR	Q7se19 neurospora
695	37	77.1	413	2	O15417_HUMAN	O15417 homo sapien
696	37	77.1	414	2	Q55LS8_CRYNE	Q55ls8 cryptococcu
697	37	77.1	414	2	Q5K8Z7_CRYNE	Q5k8z7 cryptococcu
698	37	77.1	414	2	Q8NBX7_HUMAN	Q8nbx7 homo sapien
699	37	77.1	417	2	Q6PIN9_HUMAN	Q6pin9 homo sapien
700	37	77.1	426	2	Q5YUA8_NOCFA	Q5yua8 nocardia fa
701	37	77.1	429	2	Q5BFA3_EMENI	Q5bfa3 aspergillus
702	37	77.1	430	2	Q5TQU9_ANOGA	Q5tqu9 anopheles g
703	37	77.1	435	1	RAD52_ASHGO	Q756f4 ashbya goss
704	37	77.1	445	2	Q95KJ7_MACFA	Q95kj7 macaca fasc
705	37	77.1	446	2	Q5TTT0_ANOGA	Q5ttt0 anopheles g
706	37	77.1	454	2	Q9RVH3_DEIRA	Q9rvh3 deinococcus
707	37	77.1	471	2	Q9NPM5_HUMAN	Q9npm5 homo sapien
708	37	77.1	481	2	Q82CH7_STRAW	Q82ch7 streptomyce
709	37	77.1	485	2	Q7UA85_SYNPX	Q7ua85 synechococc
710	37	77.1	487	1	EBN2_EBV	P12978 epstein-bar
711	37	77.1	487	2	Q777H1_9GAMA	Q777h1 human herpe
712	37	77.1	488	2	Q6CBY3_YARLI	Q6cby3 yarrowia li
713	37	77.1	488	2	Q95UR9_SCHAM	Q95ur9 schistocerc
714	37	77.1	488	2	Q4Q2C0_LEIMA	Q4q2c0 leishmania
715	37	77.1	489	2	Q873K2_NEUCR	Q873k2 neurospora
716	37	77.1	489	2	Q73XK9_MYCPA	Q73xk9 mycobacteri
717	37	77.1	490	2	Q69023_9GAMA	Q69023 human herpe
718	37	77.1	499	2	Q6YTE2_ORYSA	Q6yte2 oryza sativ
719	37	77.1	500	2	Q69X58_ORYSA	Q69x58

835	37	77.1	1155	2	Q8P7B3_XANCP	Q8p7b3 xanthomonas
836	37	77.1	1183	2	Q7TSH6_MOUSE	Q7tsh6 mus musculus
837	37	77.1	1191	2	Q51UB0_MAGGR	Q51ub0 magnaporthe
838	37	77.1	1200	2	Q69ZP8_MOUSE	Q69zp8 mus musculus
839	37	77.1	1201	2	Q24240_DROME	Q24240 drosophila
840	37	77.1	1209	2	Q6PFF0_MOUSE	Q6pff0 mus musculus
841	37	77.1	1210	2	Q753G0_ASHGO	Q753g0 ashbya goss
842	37	77.1	1212	2	Q8MKN2_DROME	Q8mkn2 drosophila
843	37	77.1	1216	2	Q4S2D6_TETNG	Q4s2d6 tetraodon n
844	37	77.1	1239	2	Q9V4P1_DROME	Q9v4p1 drosophila
845	37	77.1	1262	1	STNB_DROME	Q24212 drosophila
846	37	77.1	1297	2	Q9V604_DROME	Q9v604 drosophila
847	37	77.1	1325	2	Q96PW1_HUMAN	Q96pw1 homo sapien
848	37	77.1	1375	2	Q4WZV7_ASPFU	Q4wzv7 aspergillus
849	37	77.1	1391	2	Q6K8Z4_ORYSA	Q6k8z4 oryza sativ
850	37	77.1	1408	2	Q6X1Z5_9ALPH	Q6x1z5 bovine herp
851	37	77.1	1426	1	BCL9_HUMAN	O00512 homo sapien
852	37	77.1	1426	2	Q5T489_HUMAN	Q5t489 homo sapien
853	37	77.1	1434	1	SYNJ2_MOUSE	Q9d2g5 mus musculus
854	37	77.1	1464	2	Q9G782_DROME	Q9g782 drosophila
855	37	77.1	1464	2	Q9VX75_DROME	Q9vx75 drosophila
856	37	77.1	1495	2	Q5KIY9_CRYNE	Q5kiy9 cryptococcu
857	37	77.1	1496	1	SYNJ2_RAT	Q55207 rattus norv
858	37	77.1	1525	2	Q5STE9_CRYNE	Q5ste9 cryptococcu
859	37	77.1	1553	2	Q4P0W7_USTMA	Q4p0w7 ustilago ma
860	37	77.1	1560	2	Q96JP2_HUMAN	Q96jp2 homo sapien
861	37	77.1	1581	2	Q6BUZ0_DEBHA	Q6buz0 debaryomyce
862	37	77.1	1589	2	Q9UUQ9_METAN	Q9uug9 metarhizium
863	37	77.1	1661	2	Q5TQE9_ANOGA	Q5tqe9 anopheles g
864	37	77.1	1704	2	Q4I2J0_GIBZE	Q4i2j0 gibberella
865	37	77.1	1758	2	O16625_CAEEL	O16625 caenorhabdi
866	37	77.1	1790	2	Q642I4_BRARE	Q642i4 brachydanio
867	37	77.1	1884	2	Q5CS67_CRYPV	Q5cs67 cryptospori
868	37	77.1	1922	2	Q5CID8_CRYHO	Q5cid8 cryptospori
869	37	77.1	2103	2	Q5K4K9_CHLRE	Q5k4k9 chlamydomon
870	37	77.1	2375	2	Q5AN08_CANAL	Q5an08 candida alb
871	37	77.1	2383	2	Q5AMK3_CANAL	Q5amk3 candida alb
872	37	77.1	2472	1	NCOR2_MOUSE	Q9wu42 mus musculu
873	37	77.1	2551	2	Q4P0X4_USTMA	Q4p0x4 ustilago ma
874	37	77.1	2594	2	Q4SVD5_TETNG	Q4svd5 tetraodon n
875	37	77.1	2710	2	Q89U50_BRAJA	Q89u50 bradyrhizob
876	37	77.1	3122	2	P89459_HHV2	P89459 human herpe
877	37	77.1	3247	2	Q65553_9ALPH	Q65553 bovine herp
878	37	77.1	3247	2	Q77CD4_9ALPH	Q77cd4 bovine herp
879	37	77.1	3409	2	Q6SSE6_CHLRE	Q6sse6 chlamydomon
880	37	77.1	3592	2	Q7W692_BORPA	Q7w692 bordetella
881	37	77.1	3634	2	Q9JP78_BORBR	Q9jp78 bordetella
882	37	77.1	3652	2	Q7WI61_BORBR	Q7wi61 bordetella
883	37	77.1	4027	2	Q5I2R0_CHLIN	Q5i2r0 chlamydomon
884	37	77.1	5542	2	Q7YXX2_CRYPV	Q7yxx2 cryptospori
885	36.5	76.0	70	2	Q7F440_ORYSA	Q7f440 oryza sativ
886	36.5	76.0	73	2	Q5JLV0_ORYSA	Q5jlv0 oryza sativ
887	36.5	76.0	181	2	Q75LV2_ORYSA	Q75lv2 oryza sativ
888	36.5	76.0	226	1	GSPB_AERHY	P45755 aeromonas h
889	36.5	76.0	651	2	Q8S5M7_ORYSA	Q8s5m7 oryza sativ
890	36.5	76.0	890	1	TYRO3_HUMAN	Q06418 homo sapien
891	36.5	76.0	890	2	Q86VR3_HUMAN	Q86vr3 homo sapien
892	36.5	76.0	951	2	Q59FM9_HUMAN	Q59fm9 homo sapien
893	36.5	76.0	1085	2	Q5ZIE8_CHICK	Q5zie8 gallus gall
894	36.5	76.0	1307	2	Q9V4J6_DROME	Q9v4j6 drosophila
895	36	75.0	46	2	Q99LN4_MOUSE	Q99ln4 mus musculu
896	36	75.0	59	2	Q8GSH8_ORYSA	Q8gsh8 oryza sativ
897	36	75.0	60	2	Q9QVN2_9MURI	Q9qvn2 mus sp. . p
898	36	75.0	61	2	Q67UA3_ORYSA	Q67ua3 oryza sativ
899	36	75.0	64	2	Q8N507_HUMAN	Q8n507 homo sapien
900	36	75.0	66	2	Q745U5_THET2	Q745u5 thermus the
901	36	75.0	66	2	Q53WD4_THET8	Q53wd4 thermus the
902	36	75.0	68	2	Q4WAT6_ASPFU	Q4wat6 aspergillus
903	36	75.0	71	2	Q4SU74_TETNG	Q4su74 tetraodon n
904	36	75.0	72	2	Q6UU35_ORYSA	Q6uu35 oryza sativ
905	36	75.0	73	2	Q80W36_MOUSE	Q80w36 mus musculu
906	36	75.0	83	2	Q96W43_OPHNO	Q96w43 ophiostoma
907	36	75.0	86	2	Q4YHU3_PLABE	Q4yhu3 plasmodium

908	36	75.0	87	2	Q69MY3_ORYSA	Q69my3 oryza sativ
909	36	75.0	88	2	Q5Z7I4_ORYSA	Q5z7i4 oryza sativ
910	36	75.0	89	2	Q7WXL8_ALCEU	Q7wxl8 alcaligenes
911	36	75.0	93	2	Q6UU13_ORYSA	Q6uu13 oryza sativ
912	36	75.0	101	2	Q8C877_MOUSE	Q8c877 mus musculu
913	36	75.0	102	2	Q5JNC8_ORYSA	Q5jnc8 oryza sativ
914	36	75.0	103	2	Q84JC7_ORYSA	Q84jc7 oryza sativ
915	36	75.0	104	1	RK25_PEA	P11892 pisum sativ
916	36	75.0	104	2	Q9FWW1_ORYSA	Q9fww1 oryza sativ
917	36	75.0	105	2	Q9D7W2_MOUSE	Q9d7w2 mus musculu
918	36	75.0	106	2	Q9FKP0_ARATH	Q9fkp0 arabidopsis
919	36	75.0	107	2	Q8LEL5_ARATH	Q8lel5 arabidopsis
920	36	75.0	108	2	Q6A688_PROAC	Q6a688 propionibac
921	36	75.0	111	2	Q749V8_GEOSL	Q749v8 geobacter s
922	36	75.0	112	2	Q7EYJ0_ORYSA	Q7eyj0 oryza sativ
923	36	75.0	112	2	Q88EX3_PSEPK	Q88ex3 pseudomonas
924	36	75.0	114	2	Q4PEW6_USTMA	Q4pew6 ustilago ma
925	36	75.0	116	2	P82411_SPIOL	P82411 spinacia ol
926	36	75.0	116	2	Q8CDE5_MOUSE	Q8cde5 mus musculu
927	36	75.0	117	2	Q57Y88_9TRYP	Q57y88 trypanosoma
928	36	75.0	120	2	Q53783_STRAM	Q53783 streptomyce
929	36	75.0	121	2	O49510_ARATH	O49510 arabidopsis
930	36	75.0	123	2	Q8UGH2_AGRT5	Q8ugh2 agrobacteri
931	36	75.0	124	2	Q96PJ0_HUMAN	Q96pj0 homo sapien
932	36	75.0	124	2	Q7Z2T9_HUMAN	Q7z2t9 homo sapien
933	36	75.0	124	2	Q6YYD3_ORYSA	Q6yyd3 oryza sativ
934	36	75.0	126	2	Q69X78_ORYSA	Q69x78 oryza sativ
935	36	75.0	128	2	Q6QNH4_9ORTH	Q6qnh4 chorthippus
936	36	75.0	129	2	O5PQ17_XENLA	O5pq17 xenopus lae
937	36	75.0	129	2	Q4TGN5_TETNG	Q4tgn5 tetraodon n
938	36	75.0	133	2	Q950M7_9FUNG	Q950m7 rhizophydiu
939	36	75.0	135	2	Q6YSJ4_ORYSA	Q6ysj4 oryza sativ
940	36	75.0	136	2	Q8SYT0_DROME	Q8syto drosophila
941	36	75.0	136	2	Q6IIM7_DROME	Q6iim7 drosophila
942	36	75.0	138	2	Q871I5_NEUCR	Q871i5 neurospora
943	36	75.0	138	2	O5QN36_ORYSA	O5qn36 oryza sativ
944	36	75.0	139	2	Q7XI90_ORYSA	Q7xi90 oryza sativ
945	36	75.0	140	2	Q827D8_STRAW	Q827d8 streptomyce
946	36	75.0	141	2	Q4N6K2_THEPA	Q4n6k2 theileria p
947	36	75.0	141	2	Q4J3P2_AZOVI	Q4j3p2 azotobacter
948	36	75.0	144	2	Q8CCL8_MOUSE	Q8ccl8 mus musculu
949	36	75.0	147	2	Q8BFT5_MOUSE	Q8bft5 m mus muscu
950	36	75.0	147	2	Q4R4K1_MACFA	Q4r4k1 macaca fasc
951	36	75.0	147	2	O5JK95_ORYSA	O5jk95 oryza sativ
952	36	75.0	149	2	Q63IU4_BURPS	Q63iu4 burkholderi
953	36	75.0	149	2	Q4V2Y4_BURMA	Q4v2y4 burkholderi
954	36	75.0	151	2	Q6ZHK8_ORYSA	Q6zhk8 oryza sativ
955	36	75.0	151	2	Q72E50_DESVH	Q72e50 desulfovibr
956	36	75.0	152	2	Q5VXA5_HUMAN	Q5vxa5 homo sapien
957	36	75.0	152	2	Q9DUD0_9VIRU	Q9dud0 torque teno
958	36	75.0	153	2	Q17269_BRUPA	Q17269 brugia paha
959	36	75.0	153	2	Q6AGN9_LEIXX	Q6agn9 leifsonia x
960	36	75.0	154	2	Q5Z5L8_ORYSA	Q5z5l8 oryza sativ
961	36	75.0	155	2	Q6ZNT1_HUMAN	Q6znt1 homo sapien
962	36	75.0	157	2	Q59JQ4_CANAL	Q59jq4 candida alb
963	36	75.0	157	2	Q5Z770_ORYSA	Q5z770 oryza sativ
964	36	75.0	158	2	Q6Z793_ORYSA	Q6z793 oryza sativ
965	36	75.0	158	2	Q8L490_ORYSA	Q8l490 oryza sativ
966	36	75.0	158	2	Q72KK8_THET2	Q72kk8 thermus the
967	36	75.0	158	2	Q5SKI1_THET8	Q5skil thermus the
968	36	75.0	158	2	Q9D VW0_9BACU	Q9dvw0 plutella xy
969	36	75.0	159	2	Q96CB5_HUMAN	Q96cb5 homo sapien
970	36	75.0	159	2	Q9NUM6_HUMAN	Q9num6 homo sapien
971	36	75.0	159	2	Q4RY14_TETNG	Q4ry14 tetraodon n
972	36	75.0	160	2	Q9DEP6_SPAAU	Q9dep6 sparus aura
973	36	75.0	162	2	Q55IY1_CRYNE	Q55iy1 cryptococcu
974	36	75.0	162	2	Q5KCS2_CRYNE	Q5kcs2 cryptococcu
975	36	75.0	164	2	O08548_MOUSE	O08548 mus musculu
976	36	75.0	164	2	Q5J2U5_MOUSE	Q5j2u5 mus musculu
977	36	75.0	165	2	Q560J4_CRYNE	Q560j4 cryptococcu
978	36	75.0	165	2	Q96PC6_HUMAN	Q96pc6 homo sapien
979	36	75.0	165	2	Q8CI30_MOUSE	Q8ci30 mus musculu
980	36	75.0	166	2	Q6H5I4_ORYSA	Q6h5i4 oryza sativ

981	36	75.0	168	2	Q98911_CHICK	Q98911 gallus gall
982	36	75.0	169	2	Q9Y976_AERPE	Q9Y976 aeropyrum p
983	36	75.0	169	2	O35315_MOUSE	O35315 mus musculu
984	36	75.0	170	1	IR04_HCMVA	P09694 human cytom
985	36	75.0	174	2	Q9BGT7_MACFA	Q9bgt7 macaca fasc
986	36	75.0	179	2	Q94LB7_ORYSA	Q94lb7 oryza sativ
987	36	75.0	182	2	Q86F22_SCHJA	Q86f22 schistosoma
988	36	75.0	184	2	Q69K83_ORYSA	Q69k83 oryza sativ
989	36	75.0	184	2	Q8BJ04_MOUSE	Q8bj04 mus musculu
990	36	75.0	187	2	Q6Z9Z2_ORYSA	Q6z9z2 oryza sativ
991	36	75.0	189	2	Q6F6I7_HUMAN	Q6f6i7 homo sapien
992	36	75.0	189	2	Q8LGR6_VITVI	Q8lgr6 vitis vinif
993	36	75.0	192	2	Q518S5_MAGGR	Q518s5 magnaporthe
994	36	75.0	195	2	Q7QGA4_ANOGA	Q7qga4 anopheles g
995	36	75.0	195	2	Q6H4Q0_ORYSA	Q6h4q0 oryza sativ
996	36	75.0	196	2	O14557_HUMAN	O14557 homo sapien
997	36	75.0	198	2	Q8H7G8_ARATH	Q8h7g8 arabidopsis
998	36	75.0	199	1	SEC65_SCHPO	O94407 schizosacch
999	36	75.0	199	2	Q57YJ3_9TRYP	Q57yj3 trypanosoma
1000	36	75.0	199	2	Q9N3U3_CAEL	Q9n3u3 caenorhabdi

ALIGNMENTS

```
RESULT 1
Q4I257_GIBZE
ID Q4I257_GIBZE PRELIMINARY; PRT; 1112 AA.
AC Q4I257;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=FG08701.1;
OS Gibberella zeae PH-1.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
OX NCBI_TaxID=229533;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PH-1;
RA Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavkiy L.,
RA Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArellano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neill D.,
RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA Roman J., Schauer S., Schupback R., Seaman S., Severy P., Smirnov S.,
RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA Lander E.;
RT "Fusarium graminearum genome sequence.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AACM01000352; EAA70870.1; -; Genomic_DNA.
DR Hypothetical protein.
KW SEQUENCE 1112 AA; 126572 MW; 70CE0CD407F3B2C1 CRC64;
```

Query Match 100.0%; Score 48; DB 2; Length 1112;

```
Best Local Similarity 88.9%; Pred. No. 2.7e+02;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
Db 714 RPLPPLPTP 722

RESULT 2
KCNAS HUMAN
ID KCNA5 HUMAN STANDARD; PRT; 613 AA.
AC P22460;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Potassium voltage-gated channel subfamily A member 5 (Voltage-gated
DE potassium channel subunit Kv1.5) (HK2) (HPCN1).
GN Name=KCNAS;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Heart;
RX MEDLINE=91160866; PubMed=2001794;
RA Tamkun M.M., Knoth K.M., Walbridge J.A., Kroemer H., Roden D.M.,
RA Glover D.M.;
RT "Molecular cloning and characterization of two voltage-gated K+
RT channel cDNAs from human ventricle.";
RL PASEB J. 5:331-337(1991).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Insulinoma;
RX MEDLINE=91095456; PubMed=1986382;
RA Philipson L.H., Hice R.E., Schaefer K., Lamendola J., Bell G.I.,
RA Nelson D.J., Steiner D.F.;
RT "Sequence and functional expression in Xenopus oocytes of a human
RT insulinoma and islet potassium channel.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:53-57(1991).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Heart;
RX MEDLINE=92241872; PubMed=1349297;
RA Curran M.E., Landes G.M., Keating M.T.;
RT "Molecular cloning, characterization, and genomic localization of a
RT human potassium channel gene.";
RL Genomics 12:729-737(1992).
CC -1- FUNCTION: Mediates the voltage-dependent potassium ion
CC permeability of excitable membranes. Assuming opened or closed
CC conformations in response to the voltage difference across the
CC membrane, the protein forms a potassium-selective channel through
CC which potassium ions may pass in accordance with their
CC electrochemical gradient. May play a role in regulating the
CC secretion of insulin in normal pancreatic islets.
CC -1- SUBUNIT: Heterotetramer of potassium channel proteins. Interacts
CC via its PDZ-binding motif with DLG1. Forms a ternary complex with
CC DLG1 and CAV3 (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Pancreatic islets and insulinoma.
CC -1- DOMAIN: The amino terminus may be important in determining the
CC binding motif may play a role in modulation of channel activity
CC and/or targeting of the channel to specific subcellular
CC compartments (By similarity).
CC -1- DOMAIN: The segment S4 is probably the voltage-sensor and is
CC characterized by a series of positively charged amino acids at
CC every third position.
CC -1- SIMILARITY: Belongs to the potassium channel family. A (Shaker)
CC subfamily.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
```


CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

DR EMBL; M60451; AAA61276.1; -; mRNA.
DR EMBL; M55513; AAA36422.1; -; mRNA.
DR EMBL; M83254; AAA60146.1; -; mRNA.
DR PIR; A56031; A56031.
DR HSSP; P08510; 1HO2.
DR SMR; P22460; 120-218.
DR Ensembl; ENSG00000130037; Homo sapiens.
DR HGNC; HGNC:6224; KCNA5.
DR MIM; 176267; -.
DR GO; GO:0008076; C:voltage-gated potassium channel complex; TAS.
DR GO; GO:0005251; F:delayed rectifier potassium channel activity; TAS.
DR GO; GO:0006813; P:potassium ion transport; TAS.
DR InterPro; IPR000210; BTB_POZ.
DR InterPro; IPR005821; Ion_trans.
DR InterPro; IPR001622; K+channel_pore.
DR InterPro; IPR003091; K_channel.
DR InterPro; IPR003131; K_tetra.
DR InterPro; IPR004052; KV15channel.
DR InterPro; IPR003968; Kv_channel.
DR InterPro; IPR005820; M+channel_nlg.
DR InterPro; IPR003972; Shaker_channel.
DR Pfam; PF00520; Ion_trans; 1.
DR Pfam; PF02214; K_tetra; 1.
DR PRINTS; PR00169; KCHANNEL.
DR PRINTS; PR01512; KV15CHANNEL.
DR PRINTS; PR01491; KVCHANNEL.
DR PRINTS; PR01496; SHAKERCHANNEL.
DR SMART; SM00225; BTB; 1.
KW Glycoprotein; Ion transport; Ionic channel; Lipoprotein;
KW Multigene family; Palmitate; Phosphorylation; Potassium;
KW Potassium channel; Potassium transport; Repeat; Transmembrane;
KW Transport; Voltage-gated channel.
FT TRANSMEM 251 269 Segment S1.
FT TRANSMEM 325 345 Segment S2.
FT TRANSMEM 356 377 Segment S3.
FT TRANSMEM 398 419 Segment S4.
FT TRANSMEM 434 455 Segment S5.
FT TRANSMEM 495 516 Segment S6.
FT REPEAT 61 71 1.
FT REPEAT 72 82 2.
FT REGION 61 82 2 X 11 AA tandem repeat of D-[SP]-G-V-R-P-L-P-P-L-P.
FT MOTIF 480 485 Selectivity filter (By similarity).
FT MOTIF 611 613 PDZ-binding (By similarity).
FT COMPBIAS 94 99 Poly-Glu.
FT COMPBIAS 382 387 Poly-Gly.
FT MOD_RES 557 557 Phosphoserine (by PKA) (Potential).
FT LIPID 346 346 S-palmitoyl cysteine (Potential).
FT CARBOHYD 125 125 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 190 190 N-linked (GlcNAc. . .) (Potential).
FT CONFLICT 55 55 Missing (in Ref. 1).
FT CONFLICT 138 138 L -> Q (in Ref. 2).
FT CONFLICT 154 154 P -> R (in Ref. 3).
FT CONFLICT 187 188 RP -> G (in Ref. 1).
FT CONFLICT 214 214 R -> G (in Ref. 2).
FT CONFLICT 228 228 P -> V (in Ref. 2).
FT CONFLICT 282 282 L -> V (in Ref. 3).
FT CONFLICT 307 307 P -> A (in Ref. 1).
FT CONFLICT 579 613 GSCPLEKCNVKAKSNVDLRSLYALCLDTSRETDL -> QLPPREV (in Ref. 3).
SQ SEQUENCE 613 AA; 67169 MW; 6324F704F127C5BC CRC64;

Query Match 97.9%; Score 47; DB 1; Length 613;
Best Local Similarity 88.9%; Pred. No. 2e+02;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RPLPPLXP 12
|||||

Db 65 RPLPPLPDP 73
RESULT 3
Q4VAJ1_HUMAN
ID Q4VAJ1_HUMAN PRELIMINARY; PRT; 613 AA.
AC Q4VAJ1;
DT 13-SEP-2005 (TREMBLrel. 31, Created)
DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)
DE Potassium voltage-gated channel, shaker-related subfamily, member 5.
GN Name=KCNA5;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=PCR rescued clones;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Touchman J.W., Green E.D., Dickson M.C., Blakesley R.W., Grimwood J., Schmutz J., Myers R.M., Rodriguez A.C., Grimwood J., Schmutz J., Skalska U., Smalish D.E., Butterfield Y.S.N., Krzywinski M.I., Jones S.J.M., Marra M.A.;
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=PCR rescued clones;
RG NIH MGC Project;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=PCR rescued clones;
RG NIH MGC Project;
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC096358; AAH96358.1; -; mRNA.
DR EMBL; BC099665; AAH99665.1; -; mRNA.
DR EMBL; BC099666; AAH99666.1; -; mRNA.
DR EMBL; BC096357; AAH96357.1; -; mRNA.
DR InterPro; IPR000210; BTB_POZ.
DR InterPro; IPR005821; Ion_trans.
DR InterPro; IPR001622; K+channel_pore.
DR InterPro; IPR003091; K_channel.
DR InterPro; IPR003131; K_tetra.
DR InterPro; IPR004052; KV15channel.
DR InterPro; IPR003968; Kv_channel.
DR InterPro; IPR005820; M+channel_nlg.
DR InterPro; IPR003972; Shaker_channel.
DR Pfam; PF00520; Ion_trans; 1.
DR Pfam; PF02214; K_tetra; 1.
DR PRINTS; PR00169; KCHANNEL.
DR PRINTS; PR01512; KV15CHANNEL.
DR PRINTS; PR01491; KVCHANNEL.
DR PRINTS; PR01496; SHAKERCHANNEL.
DR SMART; SM00225; BTB; 1.
KW Ion transport; Ionic channel; Transmembrane; Transport.

SQ SEQUENCE 613 AA; 67228 MW; A5B02B27F8396E3D CRC64;

Query Match 97.9%; Score 47; DB 2; Length 613;
Best Local Similarity 88.9%; Pred. No. 2e+02;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RPLPPLXP 12
| | | | | | |
Db 65 RPLPPLDP 73

RESULT 4

IDENTIFIER: TNFL6_PIG STANDARD; PRT; 282 AA.

AC Q9BEA8; Q9SM04; Q9SN10; PRT; 282 AA.

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 13-SEP-2005 (Rel. 48, Last annotation update)

DE Tumor necrosis factor ligand superfamily member 6 (FAS antigen ligand)

DE [Contains: Tumor necrosis factor ligand superfamily member 6, membrane

DE form; Tumor necrosis factor ligand superfamily member 6, soluble

DE form].

GN Name=TNFSF6; Synonyms=TNF6; PRT; 282 AA.

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;

OC Sus.

OX NCBI_TaxID=9823;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RA MEDLINE=21322533; PubMed=11429161; DOI=10.1089/107999001300177493;

RA Muneta Y., Shimoji Y., Inumaru S., Mori Y.;

RT "Molecular cloning, characterization, and expression of porcine Fas

RT ligand (CD95 ligand).";

RL J. Interferon Cytokine Res. 21:305-312(2001).

RN [2]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=Guanxi bama miniature pig;

RA Zhu N., Young Y.;

RT "Molecular cloning and characterization of porcine Fas ligand cDNA.";

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

RN [3]

RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Lymphoid;

RA Tsuyuki S., Kono M., Bloom E.T.;

RT "Cloning and potential utility of porcine Fas ligand: overexpression

RT in porcine cells protects them from attack by human cytolytic cells.";

RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

RN [4]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=Landrace x Large Yorkshire white; TISSUE=Thymocyte;

RX MEDLINE=21653191; PubMed=11792426; DOI=10.1016/S0161-5890(01)00098-0;

RA Motegi-Ishiyama Y., Nakajima Y., Hoka S., Takagaki Y.;

RT "Porcine Fas-ligand gene: genomic sequence analysis and comparison

RT with human gene.";

RL Mol. Immunol. 38:581-586(2002).

CC -1- FUNCTION: Cytokine that binds to TNFRSF6/FAS, a receptor that

CC transduces the apoptotic signal into cells. May be involved in

CC cytotoxic T cell mediated apoptosis and in T cell development.

CC TNFRSF6/FAS-mediated apoptosis may have a role in the induction of

CC peripheral tolerance, in the antigen-stimulated suicide of mature

CC T cells, or both. Binding to the decoy receptor TNFRSF6B/DCR3

CC modulates its effects (By similarity).

CC -1- SUBUNIT: Homotrimer (Probable).

CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted (By

CC similarity).

CC -1- INDUCTION: By IL-18.

CC -1- PTM: The soluble form derives from the membrane form by

CC proteolytic processing (By similarity).

CC -1- SIMILARITY: Belongs to the tumor necrosis factor family.

CC -----

CC This Swiss-Prot entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use as long as its content is in no way modified and this statement is not

CC removed.

CC -----

DR EMBL; AB027297; BAB40919.1; -; mRNA.

DR EMBL; AY033634; AAK56449.1; -; mRNA.

DR EMBL; AF397407; AAK84408.1; -; mRNA.

DR EMBL; AB069764; BAB64291.1; -; Genomic DNA.

DR HSSP; P01375; 4TSV.

DR InterPro; IPR008064; Fas ligand.

DR InterPro; IPR006053; TNF_abc.

DR InterPro; IPR002961; TNF_C.

DR InterPro; IPR006052; TNF_family.

DR InterPro; IPR003636; TNF_subf.

DR PANTHER; PTHR15161; Fas_ligand; 1.

DR Pfam; PF02229; TNF; 1.

DR PRINTS; PR01681; FASLIGAND.

DR PRINTS; PR01234; TNECROSISFCT.

DR PRINTS; PR01237; TNFC.

DR ProDom; PD002012; TNF_subf; 1.

DR SMART; SM00207; TNF; 1.

DR PROSITE; PS00251; TNF_1; 1.

DR PROSITE; PS00049; TNF_2; 1.

KW Apoptosis; Cytokine; Glycoprotein; Signal-anchor; Transmembrane.

FT CHAIN 1 282 Tumor necrosis factor ligand superfamily

FT member 6, membrane form.

FT CHAIN 131 282 Tumor necrosis factor ligand superfamily

FT member 6, soluble form (By similarity).

FT TOPO_DOM 1 82 Cytoplasmic (Potential).

FT TRANSMEM 83 103 Signal-anchor for type II membrane

FT protein (Potential).

FT TOPO_DOM 104 282 Extracellular (Potential).

FT COMPBIAS 4 70 Pro-rich.

FT COMPBIAS 45 56 Poly-Pro.

FT SITE 130 131 Cleavage (By similarity).

FT CARBOHYD 185 185 N-linked (GlcNAc..) (Potential).

FT CARBOHYD 251 251 N-linked (GlcNAc..) (Potential).

FT CARBOHYD 261 261 N-linked (GlcNAc..) (Potential).

FT DISULFID 203 234 Potential.

FT CONFLICT 5 5 F -> L (in Ref. 4).

FT CONFLICT 57 57 T -> P (in Ref. 2).

SQ SEQUENCE 282 AA; 31756 MW; 6743DAA1145671FB CRC64;

Query Match 95.8%; Score 46; DB 1; Length 282;
Best Local Similarity 88.9%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RPLPPLXP 12
| | | | | | |
Db 62 RPLPPLPPP 70

RESULT 5

Q55ZA2_CRYNE PRELIMINARY; PRT; 1794 AA.

ID Q55ZA2_CRYNE PRELIMINARY; PRT; 1794 AA.

AC Q55ZA2;

DT 13-SEP-2005 (TrEMBLrel. 31, Created)

DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)

DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)

DE Hypothetical protein.

GN ORFNames=CNBA5860;

OS Cryptococcus neoformans var. neoformans B-3501A.

OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;

OC Tremellomycetidae; Tremellales; Filobasidiella.

OX NCBI_TaxID=283643;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=B-3501A;

RA Fung E., Hyman R.W., Rowley D., Bruno D., Miranda M., Fukushima M.,

RA Wickes B.L., Fu J., Davis R.W.;

RT "Cryptococcus neoformans serotype D sequencing.";

RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.

CC -1- CAUTION: The sequence shown here is derived from an

CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -!- SIMILARITY: Contains 1 PH domain.
DR EMBL; AA01000004; EAL23061.1; -; Genomic_DNA.
DR InterPro; IPR000977; DNA_ligase.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000904; Sec7.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF01369; Sec7; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00222; Sec7; 1.
DR PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN_1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
DR PROSITE; PS50190; SEC7; 1.
KW Coiled coil; Guanine-nucleotide releasing factor;
KW Hypothetical protein.
SQ SEQUENCE 1794 AA; 196453 MW; 98811FD2F194C994 CRC64;

Query Match 95.8%; Score 46; DB 2; Length 1794;
Best Local Similarity 88.9%; Pred. No. 8.3e+02;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLPP 12
DB 237 RPLPPLPPP 245

RESULT 6
Q5KNL7 CRYNE PRELIMINARY; PRT; 1811 AA.
AC Q5KNL7;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE ARF guanyl-nucleotide exchange factor, putative.
GN ORFNames=CNA06050;
OS Cryptococcus neoformans var. neoformans JEC21.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX NCBI_TaxID=214684;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=JEC21;
RA Loftus B.J., Fung E., Roncaglia P., Rowley D., Amedeo P., Bruno D.,
RA Vamathevan J., Miranda M., Anderson I.J., Fraser J.A., Allen J.E.,
RA Bosdet I.E., Brent M.R., Chiu R., Doering T.L., Donlin M.J.,
RA D'Souza C.A., Fox D.S., Grinberg V., Fu J., Fukushima M., Haas B.J.,
RA Huang J.C., Janbon G., Jones S.J.M., Krzywinski M.I., Kwon-Chung J.K.,
RA Lengeler K.B., Maiti R., Marra M.A., Marra R.E., Mathewson C.A.,
RA Mitchell T.G., Perteau M., Riggs F.R., Salzberg S.L., Shvartsbeyn A.,
RA Schein J.E., Shin H., Specht C.A., Suh B., Tenney A., Utterback T.,
RA Wickes B.L., Wye N.H., Kronstad J., Lodge J.K., Heitman J.,
RA Davis R.W., Fraser C.M., Hyman R.W.;
RT "The genome and transcriptome of Cryptococcus neoformans, a
RT basidiomycete fungal pathogen of humans."
RL Science 0:0-0(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=JEC21;
RA Loftus B., Amedeo P., Roncaglia P., Vamathevan J., Utterback T.,
RA Van Aken S., Fraser C.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=JEC21;
RX PubMed=15653466; DOI=10.1126/science.1103773;
RA Loftus B.J., Fung E., Roncaglia P., Rowley D., Amedeo P., Bruno D.,
RA Vamathevan J., Miranda M., Anderson I.J., Fraser J.A., Allen J.E.,
RA Bosdet I.E., Brent M.R., Chiu R., Doering T.L., Donlin M.J.,
RA D'Souza C.A., Fox D.S., Grinberg V., Fu J., Fukushima M., Haas B.J.,
RA Huang J.C., Janbon G., Jones S.J.M., Koo H.L., Krzywinski M.I.,
RA Kwon-Chung K.J., Lengeler K.B., Maiti R., Marra M.A., Marra R.E.,
RA Mathewson C.A., Mitchell T.G., Perteau M., Riggs F.R., Salzberg S.L.,

RA Schein J.E., Shvartsbeyn A., Shin H., Shumway M., Specht C.A.,
RA Suh B.B., Tenney A., Utterback T.R., Wickes B.L., Wortman J.R.,
RA Wye N.H., Kronstad J.W., Lodge J.K., Heitman J., Davis R.W.,
RA Fraser C.M., Hyman R.W.;
RT "The genome of the basidiomycetous yeast and human pathogen
RT Cryptococcus neoformans."
RL Science 307:1321-1324(2005).
CC -!- SIMILARITY: Contains 1 PH domain.
DR EMBL; AE017341; AAW41140.1; -; Genomic_DNA.
DR InterPro; IPR000977; DNA_ligase.
DR InterPro; IPR001849; PH.
DR InterPro; IPR011993; PH type.
DR InterPro; IPR000904; Sec7.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF01369; Sec7; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00222; Sec7; 1.
DR PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN_1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
DR PROSITE; PS50190; SEC7; 1.
KW Coiled coil; Complete proteome; Guanine-nucleotide releasing factor.
SQ SEQUENCE 1811 AA; 198207 MW; 33888A8C9C924019 CRC64;

Query Match 95.8%; Score 46; DB 2; Length 1811;
Best Local Similarity 88.9%; Pred. No. 8.4e+02;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLPP 12
DB 237 RPLPPLPPP 245

RESULT 7
FDSCP HUMAN STANDARD; PRT; 85 AA.
ID FDSCP HUMAN STANDARD; PRT; 85 AA.
AC Q8NFU4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Follicular dendritic cell secreted peptide precursor (FDC-SP) (FDC
DE secreted protein).
GN Name=FDCSP; Synonyms=C4orf7; ORFNames=UNQ733/PRO1419;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE, AND TISSUE SPECIFICITY.
RC TISSUE=Tonsil;
RX MEDLINE=22181501; PubMed=12193705;
RA Marshall A.J., Du Q., Draves K.E., Shikishima Y., Hayglass K.T.,
RA Clark E.A.;
RT "FDC-SP, a novel secreted protein expressed by follicular dendritic
RT cells."
RL J. Immunol. 169:2381-2389(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;
RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D.T., Brush J.,
RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
RA Eaton D., Foster J.S., Grimaldi C., Gu Q., Hass P.E., Heldens S.,
RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
RA Lewis L., Liao D., Mark M.R., Robbie E., Sanchez C., Schoenfeld J.,
RA Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
RA Vandlien R.L., Watanabe C., Wieand D., Woods K., Xie M.-H.,
RA Yansura D.G., Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A.D.,
RA Wood W.I., Godowski P.J., Gray A.M.;
RT "The secreted protein discovery initiative (SPDI), a large-scale
RT effort to identify novel human secreted and transmembrane proteins: a
RT bioinformatics assessment."
RL Genome Res. 13:2265-2270(2003).
RN [3]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Thyroid;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahay J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Kraywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Can bind to the surface of B lymphoma cells, but not T
lymphoma cells, consistent with a function as a secreted mediator
acting upon B cells.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Abundantly expressed in tonsil, lymph node,
and trachea; strong expression in prostate; lower expression in
thyroid, stomach, and colon.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC -----
DR EMBL; AF435080; AAN01116.1; -; mRNA.
DR EMBL; AY358958; AAQ89317.1; -; mRNA.
DR EMBL; BC062213; AAH62213.1; -; mRNA.
DR Ensembl; ENSG00000181617; Homo sapiens.
DR HGNC; HGNC:19215; C4orf7.
DR MIM; 607241; -.
KW Signal.
FT SIGNAL 1 17 Potential.
FT CHAIN 18 85 Follicular dendritic cell secreted
peptide.
FT COMPBIAS 45 82 Pro-rich.
FT SEQUENCE 85 AA; 9700 MW; 3FEC98BDE883AB07 CRC64;
Query Match 93.8%; Score 45; DB 1; Length 85;
Best Local Similarity 77.8%; Pred. No. 46;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 4 RPLPPLPXP 12
Db 51 RLPPIPPFP 59
RESULT 8
Q540F3 HUMAN PRELIMINARY; PRT; 85 AA.
AC Q540F3;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE NPC-associated peptide.
GN Name=NAP1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;

RN NUCLEOTIDE SEQUENCE.
RA Li F., Jiang W.H., Yao K.T.;
RT "Cloning and functional research of a novel human NPC-associated gene
similar to FDC-SP.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBDJ databases.
DR EMBL; AY190326; AAO41714.1; -; mRNA.
SQ SEQUENCE 85 AA; 9700 MW; 3FEC98BDE883AB07 CRC64;
Query Match 93.8%; Score 45; DB 2; Length 85;
Best Local Similarity 77.8%; Pred. No. 46;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 4 RPLPPLPXP 12
Db 51 RLPPIPPFP 59
RESULT 9
Q91NH0 9PARA PRELIMINARY; PRT; 519 AA.
AC Q91NH0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nucleocapsid protein.
OS Tioman virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Rubulavirus.
OX NCBI_TaxID=162013;
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=21235792; PubMed=11336547; DOI=10.1006/viro.2000.0882;
RA Chua K.B., Wang L.F., Lam S.K., Cramer G., Yu M., Wise T., Boyle D.,
RA Hyatt A.D., Eaton B.T.;
RT "Tioman virus, a novel paramyxovirus isolated from fruit bats in
Malaysia.";
RL Virology 283:215-229(2001).
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=22105137; PubMed=12111411; DOI=10.1007/s00705-002-0815-5;
RA Chua K.B., Wang L.F., Lam S.K., Eaton B.T.;
RT "Full length genome sequence of Tioman virus, a novel paramyxovirus in
the genus Rubulavirus isolated from fruit bats in Malaysia.";
RL Arch. Virol. 147:1323-1348(2002).
RN NUCLEOTIDE SEQUENCE.
RA Chua K.B., Wang L.-F.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBDJ databases.
DR EMBL; AF298895; AAK63199.1; -; Genomic RNA.
DR GO; GO:0019013; C:Viral nucleocapsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002021; Paramyx_ncap.
DR Pfam; PF00973; Paramyx_ncap; 1.
KW Viral nucleoprotein.
SQ SEQUENCE 519 AA; 58437 MW; F7D0437EC7F03C9A CRC64;
Query Match 93.8%; Score 45; DB 2; Length 519;
Best Local Similarity 77.8%; Pred. No. 3.1e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 4 RPLPPLPXP 12
Db 457 RLPPIPPFP 465
RESULT 10
Q6ZS83 HUMAN PRELIMINARY; PRT; 147 AA.
AC Q6ZS83;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

RA Ridler K.A., Schlessinger D., Schueler M.G., Sehra H.K.,
RA Shaw-Smith C., Shen H., Sheridan E.M., Shownkeen R., Skuce C.D.,
RA Smith M.L., Sotheran E.C., Steingrubber H.E., Steward C.A., Storey R.,
RA Swann R.M., Swarbreck D., Tabor P.E., Taudien S., Taylor T.,
RA Teague B., Thomas K., Thorpe A., Timms K., Tracey A., Trevanion S.,
RA Tromans A.C., d'Urso M., Verduzco D., Villasana D., Waldron L.,
RA Wall M., Wang Q., Warren J., Warry G.L., Wei X., West A.,
RA Whitehead S.L., Whiteley M.N., Wilkinson J.E., Willey D.L.,
RA Williams G., Williams L., Williamson A., Williamson H., Wilming L.,
RA Woodmansey R.L., Wray P.W., Yen J., Zhang J., Zhou J., Zoghbi H.,
RA Zorilla S., Buck D., Reinhardt R., Poustka A., Rosenthal A.,
RA Lehrach H., Meindl A., Minx P.J., Hillier L.W., Willard H.F.,
RA Wilson R.K., Waterston R.H., Rice C.M., Vaudin M., Coulson A.,
RA Nelson D.L., Weinstein G., Sulston J.E., Durbin R., Hubbard T.,
RA Gibbs R.A., Beck S., Rogers J., Bentley D.R.;
RT "The DNA sequence of the human X chromosome.";
RL Nature 434:325-337(2005).
RN [5]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Muscle;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- SUBCELLULAR LOCATION: Attached to the plasma membrane (Potential).
CC -!- TISSUE SPECIFICITY: Ubiquitous.
CC -!- CAUTION: Ref.5 sequences differ from that shown due to wrong
CC choices of frame.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; Y13374; CAA73804.1; -; mRNA.
DR EMBL; AF038168; AAB97361.1; -; mRNA.
DR EMBL; AF052096; AAC28618.1; -; mRNA.
DR EMBL; D88756; BAB46920.1; -; mRNA.
DR EMBL; AL136169; CAB87958.1; -; Genomic DNA.
DR EMBL; BC002385; AAH02385.2; ALT_SEQ; mRNA.
DR EMBL; BC002410; AAH02410.2; ALT_SEQ; mRNA.
DR Ensembl; ENSG00000134590; Homo sapiens.
DR HGNC; HGNC:2569; CXK1.
DR H-InvDB; HIX0017070; -.
DR MIM; 300213; -.
KW Lipoprotein; Membrane; Prenylation.
FT LIPID 206 206 S-farnesyl cysteine (Potential).
SQ SEQUENCE 209 AA; 22278 MW; 976E0D2AE93023A8 CRC64;

Query Match 89.6%; Score 43; DB 1; Length 209;
Best Local Similarity 87.5%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLPXP 12
|||||

Db 119 PLPPLPSP 126

RESULT 18
Q6IBF1 HUMAN
ID Q6IBF1_HUMAN PRELIMINARY; PRT; 209 AA.
AC Q6IBF1;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE CXK1 protein.
GN Name=CXK1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Ebert L., Schick M., Neubert P., Schatten R., Henze S., Korn B.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR456853; CAG33134.1; -; mRNA.
SQ SEQUENCE 209 AA; 22277 MW; 976E0D2AE93023A8 CRC64;

Query Match 89.6%; Score 43; DB 2; Length 209;
Best Local Similarity 87.5%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLPXP 12
|||||
Db 119 PLPPLPSP 126

RESULT 19
Q51CQ5 ENTHI
ID Q51CQ5_ENTHI PRELIMINARY; PRT; 238 AA.
AC Q51CQ5;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=19.t00018;
OS Entamoeba histolytica HM-1:IMSS.
OC Eukaryota; Entamoebidae; Entamoeba.
OX NCBI_TaxID=294381;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HM-1:IMSS;
RX PubMed=15729342; DOI=10.1038/nature03291;
RA Loftus B., Anderson I., Davies R., Alsmark U.C., Samuelson J.,
RA Amedeo P., Roncaglia P., Berriman M., Hirt R.P., Mann B.J., Nozaki T.,
RA Suh B., Pop M., Duchene M., Ackers J., Tannich E., Leippe M.,
RA Hofer M., Bruchhaus I., Willhoef U., Bhattacharya A.,
RA Chillingworth T., Churcher C., Hance Z., Harris B., Harris D.,
RA Jagels K., Moule S., Mungall K., Ormond D., Squares R., Whitehead S.,
RA Quail M.A., Rabinowitsch E., Norbertczak H., Price C., Wang Z.,
RA Guillen N., Gilchrist C., Stroup S.E., Bhattacharya S., Lohia A.,
RA Foster P.G., Sichert-Ponten T., Weber C., Singh U., Mukherjee C.,
RA El-Sayed N.M., Petri W.A., Clark C.G., Embley T.M., Barrell B.,
RA Fraser C.M., Hall N.;
RT "The genome of the protist parasite Entamoeba histolytica.";
RL Nature 433:865-868(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAFB0100091; EAL50578.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 238 AA; 27384 MW; 3EA13C9FC5259A8B CRC64;

Query Match 89.6%; Score 43; DB 2; Length 238;
Best Local Similarity 77.8%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
| | | | |
Db 124 RLPPTTP 132

RESULT 20
Q4QHFL1 LEIMA
ID Q4QHFL1 LEIMA PRELIMINARY; PRT; 243 AA.
AC Q4QHFL1
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=LmjF10.0640;
OS Leishmania major.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Friedlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neil S.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; CT005249; CAJ02703.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 243 AA; 25786 MW; 9CACE2038EDF767E CRC64;

Query Match 89.6%; Score 43; DB 2; Length 243;
Best Local Similarity 87.5%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLXP 12
| | | | |
Db 151 PLPPLSP 158

RESULT 21
UDU3 ARATH
ID UDU3 ARATH STANDARD; PRT; 287 AA.
AC Q9FHD3;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical DUF26 domain containing protein 3 precursor.
GN OrderedLocusNames=At5g41300; ORFNames=K1013.10;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=cv. Columbia;
RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty p1 and TAC
RT clones.";
RL DNA Res. 7:31-63(2000).
CC -|- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
CC (potential).
CC -|- SIMILARITY: Contains 2 DUF26 domains.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AB019225; BAB1106.1; -; Genomic_DNA.
DR GeneFarm; 704; 92.

DR InterPro; IPR002902; DUF26.
DR Pfam; PF01657; DUF26; 2.
KW Glycoprotein; GPI-anchor; Hypothetical protein; Lipoprotein; Membrane;
KW Repeat; Signal.
FT SIGNAL 1 26 Potential.
FT CHAIN 27 269 Hypothetical DUF26 domain containing
FT PROPEP 270 287 protein 3.
FT DOMAIN 82 136 Removed in mature form (Potential).
FT DOMAIN 198 251 DUF26 1.
FT LIPID 269 269 DUF26 2.
FT CARBOHYD 43 43 GPI-anchor amidated glycine (Potential).
FT CARBOHYD 47 47 N-linked (GlcNAc. .) (Potential).
FT CARBOHYD 63 63 N-linked (GlcNAc. .) (Potential).
FT CARBOHYD 72 72 N-linked (GlcNAc. .) (Potential).
FT CARBOHYD 93 93 N-linked (GlcNAc. .) (Potential).
FT CARBOHYD 103 103 N-linked (GlcNAc. .) (Potential).
FT CARBOHYD 111 111 N-linked (GlcNAc. .) (Potential).
FT CARBOHYD 212 212 N-linked (GlcNAc. .) (Potential).
SQ SEQUENCE 287 AA; 32475 MW; B429DCDF9BAB1B8F CRC64;

Query Match 89.6%; Score 43; DB 1; Length 287;
Best Local Similarity 87.5%; Pred. No. 3.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLXP 12
| | | | |
Db 251 PLPPLSP 258

RESULT 22
Q9FB15 9ACTO
ID Q9FB15 9ACTO PRELIMINARY; PRT; 462 AA.
AC Q9FB15;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Peptide synthetase NRPS.
GN Name=blmII;
OS Streptomyces verticillus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=29309;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC15003;
RX MEDLINE=20389599; PubMed=10930733;
RA Du L., Chen M., Sanchez C., Shen B.;
RT "An oxidation domain in the BlmIII non-ribosomal peptide synthetase
RT probably catalyzing thiazole formation in the biosynthesis of the
RT anti-tumor drug bleomycin in Streptomyces verticillus ATCC15003.";
RL FEMS Microbiol. Lett. 189:171-175(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC15003;
RX MEDLINE=20500448; PubMed=11048953; DOI=10.1016/S1074-5521(00)00011-9;
RA Du L., Sanchez C., Chen M., Edwards D.J., Shen B.;
RT "The biosynthetic gene cluster for the antitumor drug bleomycin from
RT Streptomyces verticillus ATCC15003 supporting functional interactions
RT between nonribosomal peptide synthetases and a polyketide synthase.";
RL Chem. Biol. 7:623-642(2000).
DR EMBL; AF210249; AAG02367.1; -; Genomic_DNA.
DR InterPro; IPR001242; Condensatn.
DR Pfam; PF00668; Condensation; 1.
SQ SEQUENCE 462 AA; 49808 MW; 11593C8433F96D00 CRC64;

Query Match 89.6%; Score 43; DB 2; Length 462;
Best Local Similarity 87.5%; Pred. No. 5.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLXP 12
| | | | |
Db 447 PLPPLPAP 454


```
RESULT 23
Q9KW38_9RICK PRELIMINARY; PRT; 486 AA.
AC Q9KW38;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Component of type IV secretion system.
GN Name=virB10;
OS Wolbachia sp. wKueYO.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Wolbachieae; Wolbachia.
OX NCBI_TaxID=130485;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=WKueYO;
RA Masui S., Sasaki T., Ishikawa H.;
RT "Genes for the Type IV Secretion System in an Intracellular Symbiont,
RT Wolbachia, a Causative Agent of Various Sexual Alterations in
RT Arthropods.";
RL J. Bacteriol. 182:6529-6531(2000).
DR EMBL; AB045235; BAA97441.1; -; Genomic_DNA.
DR GO; GO:0009291; P:unidirectional conjugation; IEA.
DR InterPro; IPR005498; TrbI.
DR Pfam; PF03743; TrbI; 1.
SQ SEQUENCE 486 AA; 53026 MW; 590D0DE79240DEBF CRC64;

Query Match 89.6%; Score 43; DB 2; Length 486;
Best Local Similarity 87.5%; Pred. No. 5.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLXP 12
Db 95 PLPPLTP 102

RESULT 24
Q73IY6_WOLPM PRELIMINARY; PRT; 486 AA.
AC Q73IY6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Type IV secretion system protein VirB10.
GN OrderedLocusNames=WD0006;
OS Wolbachia pipientis wMel.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Wolbachieae; Wolbachia.
OX NCBI_TaxID=66077;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15024419; DOI=10.1371/journal.pbio.0020069;
RA Wu M., Sun L.V., Vamathevan J.J., Riegler M., DeBoy R.T.,
RA Brownlie J.C., McGraw E.A., Martin W., Esser C., Ahmadinejad N.,
RA Wiegand C., Madupu R., Beanan M.J., Brinkac L.M., Daugherty S.C.,
RA Durkin A.S., Kolonay J.F., Nelson W.C., Mohamoud Y., Lee P.,
RA Berry K.J., Young M.B., Utterback T.R., Weidman J.F., Nierman W.C.,
RA Paulsen I.T., Nelson K.B., Tettelin H., O'Neill S.L., Eisen J.A.;
RT "Phylogenomics of the reproductive parasite Wolbachia pipientis wMel:
RT a streamlined genome overrun by mobile genetic elements.";
RL PLoS Biol. 2:327-341(2004).
DR EMBL; AE017256; AAS13774.1; -; Genomic_DNA.
DR TIGR; WD0006; -.
DR GO; GO:0009291; P:unidirectional conjugation; IEA.
DR InterPro; IPR005498; TrbI.
DR Pfam; PF03743; TrbI; 1.
KW Complete proteome.
SQ SEQUENCE 486 AA; 53012 MW; 5913FC000F69B26F CRC64;

Query Match 89.6%; Score 43; DB 2; Length 486;
Best Local Similarity 87.5%; Pred. No. 5.3e+02;
```

```
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLXP 12
Db 95 PLPPLTP 102

RESULT 25
Q63IB6_BURPS PRELIMINARY; PRT; 487 AA.
ID Q63IB6_BURPS PRELIMINARY;
AC Q63IB6;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative MFS family transporter.
GN OrderedLocusNames=BPS2155;
OS Burkholderia pseudomallei (Pseudomonas pseudomallei).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia; pseudomallei group.
OX NCBI_TaxID=28450;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K96243;
RX PubMed=15377794; DOI=10.1073/pnas.0403302101;
RA Holden M.T.G., Titball R.W., Peacock S.J., Cerdeno-Tarraga A.-M.,
RA Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.L.,
RA Bentley S.D., Sebahia M., Thomson N.R., Bason N., Beacham I.R.,
RA Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,
RA Chillingworth T., Cronin A., Crossett B., Davis P., Deshazer D.,
RA Feltwell T., Fraser A., Hance Z., Hauser H., Holroyd S., Jagels K.,
RA Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,
RA Rabinowitsch E., Rutherford K., Sanders M., Simmonds M.,
RA Songsivilai S., Stevens K., Tumapa S., Vesaratchavest M.,
RA Whitehead S., Yeats C., Barrell B.G., Oyston P.C.F., Parkhill J.;
RT "Genomic plasticity of the causative agent of melioidosis,
RT Burkholderia pseudomallei.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245(2004).
DR EMBL; BX571966; CAH39639.1; -; Genomic_DNA.
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR011701; MFS_1.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF07690; MFS_1; 1.
DR PROSITE; PS50850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 487 AA; 51782 MW; E2F6F509629CDB0C CRC64;

Query Match 89.6%; Score 43; DB 2; Length 487;
Best Local Similarity 77.8%; Pred. No. 5.3e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
Db 244 RPLPPVPPP 252

RESULT 26
Q4P068_USTMA PRELIMINARY; PRT; 488 AA.
ID Q4P068_USTMA PRELIMINARY;
AC Q4P068;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=UM06495.1;
OS Ustilago maydis 521.
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OX NCBI_TaxID=237631;
```

```
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=521;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
RA Bayul T., Blitshsteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA Gnirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamvysseilis M., Karlsson E.,
RA Kells C., Kieu A., Kisaner P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., Mcdonough S., Mcghee T., Meldrim J., Meneus L.,
RA Mesirov J., Mihalev A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., News C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
RA O'neill K., Osman S., Parker S., Perrin D., Phunkhang P., Piquani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thoulutsang Y., Topham K.,
RA Towey S., Tsamla T., Taomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Ustilago maydis.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBSJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBSJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACP01000259; EAK83544.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 488 AA; 53032 MW; 18D4BD384687A61E CRC64;

Query Match 89.6%; Score 43; DB 2; Length 488;
Best Local Similarity 87.5%; Pred. No. 5.4e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLPXP 12
Db 191 PLPPLPSP 198

RESULT 27
Q52SK2 9RICK PRELIMINARY; PRT; 493 AA.
AC Q52SK2;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE VirB10.
GN Name=virB10;
OS Wolbachia endosymbiont of Armadillidium vulgare.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Wolbachieae; Wolbachia.
OX NCBI_TaxID=77039;
RN [1]
```

```
RP NUCLEOTIDE SEQUENCE.
RA Felix C., Greve P., Braquart-Varnier C., Braig H.R., Martin G.;
RT "Characterization and transcriptional analysis of two gene clusters
RT for a type IV secretion machinery in feminizing intracellular symbiont
RT Wolbachia.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AY967767; AAX86707.1; -; Genomic DNA.
SQ SEQUENCE 493 AA; 54029 MW; 3B889957279FA02D CRC64;

Query Match 89.6%; Score 43; DB 2; Length 493;
Best Local Similarity 87.5%; Pred. No. 5.4e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLPXP 12
Db 95 PLPPLPTP 102

RESULT 28
Q4WXJ5 ASPFU PRELIMINARY; PRT; 551 AA.
ID Q4WXJ5 ASPFU PRELIMINARY; PRT; 551 AA.
AC Q4WXJ5;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=Afu3g09730;
OS Aspergillus fumigatus Af293.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=330879;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Af293;
RA Nierman W., Pain A., Anderson M.J., Wortman J., Kim H. Stanley.,
RA Arroya J., Berriman M., Abe K., Archer D.B., Bermejo C., Bennett J.,
RA Bowyer P., Chen D., Collins M., Coulson R., Davies R., Dyer P.S.,
RA Farman M., Fedorova N., Fedorova N., Feldblyum T.V., Fischer R.,
RA Fosker N., Fraser A., Garcia J.L., Garcia M.J., Goble A.,
RA Goldman G.H., Gomi K., Griffith-Jones S., Gwilliam R., Haas B.,
RA Haas H., Harris D., Horiuchi H., Huang J., Humphrey S., Jimenez J.,
RA Keller N., Khouri H., Kitamoto K., Kobayashi T., Kulkarni R.,
RA Kumagai T., Lafton A., Latge J.-P., Li W., Lord A., Lu C.,
RA Majoros W.H., May G.S., Miller B.L., Mohamoud Y., Molina M., Monod M.,
RA Mouyna I., Mulligan S., Murphy L., O'Neil S., Paulsen I.,
RA Penalva M.A., Pertea M., Price C., Pritchard B.L., Quail M.A.,
RA Rabinowitsch E., Rawlins N., Rajandream M.-A., Reichard U.,
RA Renauld H., Robson G.D., Rodriguez de Cordoba S., Rodriguez-Pena J.M.,
RA Ronning C.M., Rutter S., Salzberg S.L., Sanchez M.,
RA Sanchez-Ferrero J.C., Saunders D., Seeger K., Squares R., Squares S.,
RA Takeuchi M., Tekala F., Turner G., Vazquez de Aldana C.R., Weidman J.,
RA White O., Woodward J., Yu J.-H., Fraser C., Galagan J.E., Asai K.,
RA Machida M., Hall N., Barrell B., Denning D.W.;
RT "Genomic sequence of the pathogenic and allergenic filamentous fungus
RT Aspergillus fumigatus.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBSJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBSJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHF01000002; EAL92608.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 551 AA; 61426 MW; 6743CC767983CFD3 CRC64;

Query Match 89.6%; Score 43; DB 2; Length 551;
Best Local Similarity 77.8%; Pred. No. 6.1e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLPXP 12
Db 242 QPLPPLPNP 250

RESULT 29
```

Q627E0 ORYSA
ID Q627E0_ORYSA PRELIMINARY; PRT; 554 AA.
AC Q627E0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative pentatricopeptide (PPR) repeat-containing protein.
GN Name=P0020C11.29-1;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, PAC
clone:P0020C11.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP004865; BADI5828.1; -; Genomic_DNA.
DR Gramene; Q627E0; -;
DR InterPro; IPR002865; PPR.
DR InterPro; IPR011990; TPR-like_helical.
DR Pfam; PF01535; PPR; 13.
DR TIGRFAMs; TIGR00756; PPR; 12.
KW Repeat.
SQ SEQUENCE 554 AA; 60020 MW; 440B2408B8CFE044 CRC64;

Query Match 89.6%; Score 43; DB 2; Length 554;
Best Local Similarity 77.8%; Pred. No. 6.1e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLPXP 12
|||
Db 26 RPYPPPLPAP 34

RESULT 30
Q55P29 CRYNE
ID Q55P29_CRYNE PRELIMINARY; PRT; 640 AA.
AC Q55P29;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=CNBG3830;
OS Cryptococcus neoformans var. neoformans B-3501A.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX NCBI_TaxID=283643;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B-3501A;
RA Fung E., Hyman R.W., Rowley D., Bruno D., Miranda M., Fukushima M.,
Wickes B.L., Fu J., Davis R.W.;
RT "Cryptococcus neoformans serotype D sequencing.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AA01000038; EAL19755.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 640 AA; 67879 MW; 63F41A83A7E7E1ED CRC64;

Query Match 89.6%; Score 43; DB 2; Length 640;
Best Local Similarity 87.5%; Pred. No. 7.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLPXP 12
|||
Db 130 PLPPLPSP 137

RESULT 31
Q6ZON5 ORYSA
ID Q6ZON5_ORYSA PRELIMINARY; PRT; 745 AA.
AC Q6ZON5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Putative threonyl-tRNA synthetase.
GN Name=OJ1212_C09.6; Synonyms=OJ1705_A03.38;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Katayose Y.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005487; BAD01414.1; -; Genomic DNA.
DR EMBL; AP003918; BAD01219.1; -; Genomic_DNA.
DR Gramene; Q6ZON5; -;
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0004829; F:threonine-tRNA ligase activity; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR GO; GO:0006435; P:threonyl-tRNA aminoacylation; IEA.
DR InterPro; IPR004154; HGTP_anticonodon.
DR InterPro; IPR004095; TGS.
DR InterPro; IPR02314; tRNA-synt_2b.
DR InterPro; IPR002320; tRNA-synt_thr.
DR InterPro; IPR006195; tRNA_ligase_II.
DR Pfam; PF03129; HGTP_anticonodon; 1.
DR Pfam; PF02824; TGS; 1.
DR Pfam; PF00587; tRNA-synt_2b; 1.
DR PRINTS; PR01047; TRNASYNTHTHR.
DR TIGRFAMs; TIGR00418; thrS; 1.
DR PROSITE; PS50862; AA_TRNA_LIGASE_II; 1.
KW Aminoacyl-tRNA synthetase.
SQ SEQUENCE 745 AA; 84945 MW; BFD0C3D44974465 CRC64;

Query Match 89.6%; Score 43; DB 2; Length 745;
Best Local Similarity 77.8%; Pred. No. 8.3e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLPXP 12
|||
Db 19 RPFPPPLPRP 27

RESULT 32
Q4T7J1 TETNG
ID Q4T7J1_TETNG PRELIMINARY; PRT; 851 AA.
AC Q4T7J1;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome undetermined SCAF8089, whole genome shotgun sequence.
GN ORFNames=GSTENG00005703001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,

RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castellì V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff JN., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RL the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAE01008089; CAF91141.1; -; Genomic DNA.
SQ SEQUENCE 851 AA; 94830 MW; A952ACD44D3388F4 CRC64;

Query Match 89.6%; Score 43; DB 2; Length 851;
Best Local Similarity 77.8%; Pred. No. 9.6e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
:|||||
Db 286 KPLPPLPPP 294

RESULT 33
Q4SDM9 TETNG
ID Q4SDM9_TETNG PRELIMINARY; PRT; 963 AA.
AC Q4SDM9;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 10 SCAF14634, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00019959001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castellì V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff JN., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RL the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -!- SIMILARITY: Contains 1 C2 domain.
DR EMBL; CAAE01014634; CAG01253.1; -; Genomic_DNA.

DR InterPro; IPR000008; C2.
DR Pfam; PF00168; C2; 1.
DR PRINTS; PR00360; C2DOMAIN.
DR SMART; SM00239; C2; 1.
DR PROSITE; PS50004; C2 DOMAIN; 1.
FT NON TER 963
SQ SEQUENCE 963 AA; 104747 MW; F496F320D2668851 CRC64;

Query Match 89.6%; Score 43; DB 2; Length 963;
Best Local Similarity 77.8%; Pred. No. 1.1e+03;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
|||||
Db 549 RPLPPTTP 557

RESULT 34
Q7RYU1 NEUCR
ID Q7RYU1_NEUCR PRELIMINARY; PRT; 969 AA.
AC Q7RYU1;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Predicted protein.
GN Name=NCU00367.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA Kamal M., Kamvysselis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
RA Krystofova S., Rasmussen C., Metzberg R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmani S.A.,
RA DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.;
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL Nature 0:0-0(2003).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABX01000723; EAA28017.1; -; Genomic DNA.
SQ SEQUENCE 969 AA; 109976 MW; 77860EC9134F1429 CRC64;

Query Match 89.6%; Score 43; DB 2; Length 969;
Best Local Similarity 87.5%; Pred. No. 1.1e+03;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLXP 12
|||||
Db 574 PLPPLTP 581

RESULT 35
Q9U5X0 DROME
ID Q9U5X0_DROME PRELIMINARY; PRT; 1299 AA.
AC Q9U5X0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Prickle sp1e isoform.
GN Name=pk; ORFNames=CG12830;

OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=isogenic dp cn bw;
RX MEDLINE=99415814; PubMed=10485852; DOI=10.1101/gad.13.17.2315;
RA Gubb D., Green C., Huen D., Coulson D., Johnson G., Tree D.,
RA Collier S., Roote J.;
RT "The balance between isoforms of the prickly LIM domain protein is
RT critical for planar polarity in Drosophila imaginal discs.";
RL Genes Dev. 13:2315-2327(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=isogenic dp cn bw;
RA Gubb D.C.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 2 LIM zinc-binding domains.
DR EMBL; AJ243710; CAB57345.3; -; mRNA.
DR HSSP; P04006; LIML.
DR FlyBase; FBgn0003090; CG12830.
DR FlyBase; FBgn0003090; pk.
DR GO; GO:0005737; C:cytoplasm; TAS.
DR GO; GO:0016020; C:membrane; NAS.
DR GO; GO:0005515; F:protein binding; NAS.
DR GO; GO:0042067; P:establishment of ommatidial polarity (sensu. . ; IMP.
DR GO; GO:0045184; P:establishment of protein localization; TAS.
DR GO; GO:0001737; P:establishment of wing hair orientation; TAS.
DR GO; GO:0045185; P:maintenance of protein localization; TAS.
DR InterPro; IPR001781; LIM.
DR InterPro; IPR010442; PET.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00412; LIM; 2.
DR Pfam; PF06297; PET; 1.
DR SMART; SM00132; LIM; 3.
DR PROSITE; PS00478; LIM_DOMAIN_1; 2.
DR PROSITE; PS00023; LIM_DOMAIN_2; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
KW LIM domain, Metal-binding; Zinc.
SQ SEQUENCE 1299 AA; 140529 MW; 3D6D3A31717B7DE CRC64;

Query Match 89.6%; Score 43; DB 2; Length 1299;
Best Local Similarity 77.8%; Pred. No. 1.5e+03;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPPLPXP 12
Db 262 RPVPPLPPP 270

RESULT 36
Q9V4I9 DROME
ID Q9V4I9_DROME PRELIMINARY; PRT; 1299 AA.
AC Q9V4I9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CG11084-PC, isoform C.
GN Name=pk; ORFNames=CG11084;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA

RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abiril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foslter C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Kimmel B.E., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler P., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
RT melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomics perspective.";
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RG Berkeley Drosophila Genome Project;
RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,

Yu C., Rubin G.;
RT "Drosophila melanogaster release 4 sequence.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE.
RG FlyBase;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 2 LIM zinc-binding domains.
DR EMBL; AE003842; AAF59281.2; -; Genomic_DNA.
DR HSSP; P04006; IIML.
DR Ensembl; CG11084; Drosophila melanogaster.
DR FlyBase; FBgn0003090; CG12830.
DR FlyBase; FBgn0003090; pk.
DR GO; GO:0005737; C:cytoplasm; TAS.
DR GO; GO:0016020; C:membrane; NAS.
DR GO; GO:0005515; F:protein binding; NAS.
DR GO; GO:0042067; P:establishment of ommatidial polarity (sensu. . .; IMP.
DR GO; GO:0045184; P:establishment of protein localization; TAS.
DR GO; GO:0001737; P:establishment of wing hair orientation; TAS.
DR GO; GO:0045185; P:maintenance of protein localization; TAS.
DR InterPro; IPR001781; LIM.
DR InterPro; IPR010442; PET.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00412; LIM; 2.
DR Pfam; PF06297; PET; 1.
DR ProDom; PD000094; LIM; 3.
DR SMART; SM00132; LIM; 3.
DR PROSITE; PS00478; LIM_DOMAIN_1; 2.
DR PROSITE; PS00023; LIM_DOMAIN_2; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
KW LIM domain; Metal-binding; Zinc.
SQ SEQUENCE 1299 AA; 140721 MW; 8BFAF1F75F352485 CRC64;

Query Match 89.6%; Score 43; DB 2; Length 1299;
Best Local Similarity 77.8%; Pred. No. 1.5e+03;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPLPLPXP 12
||:|||||
Db 262 RPVPPLPPP 270

RESULT 37
Q4QEP3 LEIMA
ID Q4QEP3_LEIMA PRELIMINARY; PRT; 2145 AA.
AC Q4QEP3;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=LmjF16.1240;
OS Leishmania major.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Friedlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neill S.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; CT005255; CAJ03948.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 2145 AA; 227134 MW; EC538A1C8C08EF69 CRC64;

Query Match 89.6%; Score 43; DB 2; Length 2145;
Best Local Similarity 87.5%; Pred. No. 2.5e+03;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLPXP 12
|||||||
Db 1231 PLPPLPSP 1238

RESULT 38
Q5XMJ3 9PICO
ID Q5XMJ3_9PICO PRELIMINARY; PRT; 2432 AA.
AC Q5XMJ3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Polyprotein.
OS Aichi virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Kobuvirus.
OX NCBI_TaxID=72149;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BAY/1/03/DEU;
RA Oh D.-Y., Hauroeder B., Schreier E.;
RT "Isolation and characterization of Aichi virus in Germany.";
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY747174; AAV28657.1; -; Genomic_RNA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003724; F:RNA helicase activity; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR003959; AAA_ATPase_cent.
DR InterPro; IPR000199; Pept_C3_picorn.
DR InterPro; IPR004004; Pept_Calici.
DR InterPro; IPR001676; Rhv.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_P5vir.
DR Pfam; PF00004; AAA; 1.
DR Pfam; PF00548; Peptidase_C3; 1.
DR Pfam; PF00680; RdRP_1; 1.
DR Pfam; PF00073; Rhv; 2.
DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PR00918; CALICVIRUSNS.
KW Polyprotein.
SQ SEQUENCE 2432 AA; 265525 MW; BE9102F34E8979E7 CRC64;

Query Match 89.6%; Score 43; DB 2; Length 2432;
Best Local Similarity 87.5%; Pred. No. 2.9e+03;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLPXP 12
|||||||
Db 996 PLPPLPTP 1003

RESULT 39
Q91QP4 9PICO
ID Q91QP4_9PICO PRELIMINARY; PRT; 2432 AA.
AC Q91QP4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Aichi virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Kobuvirus.
OX NCBI_TaxID=72149;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21376426; PubMed=11483747;
RX DOI=10.1128/JVI.75.17.8021-8030.2001;
RA Sasaki J., Kusuhara Y., Maeno Y., Kobayashi N., Yamashita T.,
RA Sakae K., Takeda N., Taniguchi K.;
RT "Construction of an infectious cDNA clone of Aichi virus (a new member

RT of the family Picornaviridae) and mutational analysis of a stem-loop
RT structure at the 5' end of the genome.";
RL J. Virol. 75:8021-8030(2001).
DR EMBL; AB040749; BAB62889.1; -; Genomic_RNA.
DR HSSP; P03300; 1RDR.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003724; F:RNA helicase activity; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR003959; AAA ATPase_centr.
DR InterPro; IPR000199; Pept_C3_picorn.
DR InterPro; IPR004004; Pept_Calici.
DR InterPro; IPR001676; Rhv.
DR InterPro; IPR000605; RNA helicase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_Psvir.
DR Pfam; PF00004; AAA; 1.
DR Pfam; PF00548; Peptidase_C3; 1.
DR Pfam; PF00680; RdRP_1; 1.
DR Pfam; PF00073; Rhv; 2.
DR Pfam; PF00910; RNA helicase; 1.
DR PRINTS; PR00918; CALICVIRUSNS.
KW Polyprotein.
SQ SEQUENCE 2432 AA; 265456 MW; B451BABAD3CD1460 CRC64;

Query Match 89.6%; Score 43; DB 2; Length 2432;
Best Local Similarity 87.5%; Pred.No. 2.9e+03;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLXP 12
| | | | | |
Db 996 PLPPLPT 1003

RESULT 40
O91464_9PICO
ID O91464_9PICO PRELIMINARY; PRT; 2433 AA.
AC O91464;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Aichi virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Kobuvirus.
OX NCBI_TaxID=72149;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=A846/88;
RX MEDLINE=98406258; PubMed=9733894;
RA Yamashita T., Sakae K., Tsuzuki H., Suzuki Y., Ishikawa N., Takeda N.,
RA Miyamura T., Yamazaki S.;
RT "Complete nucleotide sequence and genetic organization of Aichi virus,
RT a distinct member of the Picornaviridae associated with acute
RT gastroenteritis in humans.";
RL J. Virol. 72:8408-8412(1998).
DR EMBL; AB010145; BAA31356.1; -; Genomic_RNA.
DR HSSP; P03300; 1RDR.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003724; F:RNA helicase activity; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR003959; AAA ATPase_centr.
DR InterPro; IPR000199; Pept_C3_picorn.

DR InterPro; IPR004004; Pept_Calici.
DR InterPro; IPR001676; Rhv.
DR InterPro; IPR000605; RNA helicase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_Psvir.
DR Pfam; PF00004; AAA; 1.
DR Pfam; PF00548; Peptidase_C3; 1.
DR Pfam; PF00680; RdRP_1; 1.
DR Pfam; PF00073; Rhv; 2.
DR Pfam; PF00910; RNA helicase; 1.
DR PRINTS; PR00918; CALICVIRUSNS.
KW Polyprotein.
FT CHAIN 171 540 VP0.
FT CHAIN 541 764 VP3.
FT CHAIN 765 1042 VP1.
FT CHAIN 1043 1153 2A.
FT CHAIN 1154 1318 2B.
FT CHAIN 1319 1653 2C.
FT CHAIN 1654 1748 3A.
FT CHAIN 1749 1775 3B.
FT CHAIN 1776 1965 3C.
FT CHAIN 1966 2433 3D.
SQ SEQUENCE 2433 AA; 265526 MW; 44EFBBE1B76D6188 CRC64;

Query Match 89.6%; Score 43; DB 2; Length 2433;
Best Local Similarity 87.5%; Pred.No. 2.9e+03;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLXP 12
| | | | | |
Db 997 PLPPLPT 1004

RESULT 41
Q4TWI0_9PICO
ID Q4TWI0_9PICO PRELIMINARY; PRT; 2433 AA.
AC Q4TWI0;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Polyprotein.
OS Aichi virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Kobuvirus.
OX NCBI_TaxID=72149;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Goiania/GO/03/01/Brazil;
RA Silva P.A., Cardoso D.D.P., Schreier E.;
RT "Molecular characterization of aichi virus from Brazil.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ028632; AAY46271.1; -; Genomic_RNA.
KW Polyprotein.
SQ SEQUENCE 2433 AA; 266277 MW; 298F73C2A60AE53C CRC64;

Query Match 89.6%; Score 43; DB 2; Length 2433;
Best Local Similarity 87.5%; Pred.No. 2.9e+03;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLXP 12
| | | | | |
Db 997 PLPPLPT 1004

RESULT 42
Q5WLC4_BACSK
ID Q5WLC4_BACSK PRELIMINARY; PRT; 80 AA.
AC Q5WLC4;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Inner spore coat protein D.

```
GN OrderedLocusNames=ABC0289;
OS Bacillus clausii (strain KSM-K16).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=56692;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KSM-K16;
RA Takaki Y., Kageyama Y., Shimamura S., Suzuki H., Nishi S., Hatada Y.,
RA Kawai S., Ito S., Horikoshi K.;
RA "The complete genome sequence of the alkaliphilic Bacillus clausii
RT KSM-K16.";
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP006627; BAD62831.1; -; Genomic_DNA.
KW Capsid protein; Complete proteome.
SQ SEQUENCE 80 AA; 9185 MW; 90E297ED8E2862E3 CRC64;

Query Match      87.5%; Score 42; DB 2; Length 80;
Best Local Similarity 77.8%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RLPPLPXP 12
   |||||
Db 15 RLPPLPXP 23

RESULT 43
Q4H8A2_9DEIO
ID Q4H8A2_9DEIO PRELIMINARY; PRT; 106 AA.
AC Q4H8A2;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=DgeoDRAFT_0541;
OS Deinococcus geothermalis DSM 11300.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=319795;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DSM 11300;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hammon N., Israni S., Pitluck S., Richardson P.;
RA "Sequencing of the draft genome assembly of Deinococcus geothermalis
RT DSM 11300.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DSM 11300;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RA "Annotation of the draft genome assembly of Deinococcus geothermalis
RT DSM 11300.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHE01000007; EAL82441.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 106 AA; 11553 MW; 0C49351FA56D3310 CRC64;

Query Match      87.5%; Score 42; DB 2; Length 106;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLPXP 12
   |||||
Db 8 PLPPLPXP 15

RESULT 44
Q5AGI6_CANAL
```

```
ID Q5AGI6_CANAL PRELIMINARY; PRT; 134 AA.
AC Q5AGI6;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein.
GN ORFNames=CaO19.11804;
OS Candida albicans SC5314.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=237561;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SC5314;
RX PubMed=15123810; DOI=10.1073/pnas.0401648101;
RA Jones T., Federspiel N.A., Chibana H., Dungan J., Kalman S.,
RA Magee B.B., Newport G., Thorstenson Y.R., Agabian N., Magee P.T.,
RA Davis R.W., Scherer S.;
RA "The diploid genome sequence of Candida albicans.";
RT Proc. Natl. Acad. Sci. U.S.A. 101:7329-7334(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SC5314;
RA Dungan J., Kuo A., Newport G., Lan C.-Y., Iijima C., Adegbola O.,
RA Roberts J., Persson K., Donnelly S., Favoreto S., Tzung K.-W.,
RA Jones T., Scherer S., Agabian N.;
RA "Annotation of the Genome of Candida albicans.";
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACQ01000021; EAL01934.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 134 AA; 14368 MW; 06CBC8DD0EF305BB CRC64;

Query Match      87.5%; Score 42; DB 2; Length 134;
Best Local Similarity 87.5%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLPXP 12
   |||||
Db 89 PLPPLPIP 96

RESULT 45
Q5Z6V2_ORYSA
ID Q5Z6V2_ORYSA PRELIMINARY; PRT; 144 AA.
AC Q5Z6V2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein P0551A03.21.
GN Name=P0551A03.21;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, PAC
RT clone:P0551A03.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP004818; BAD54317.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 144 AA; 15117 MW; 008F8B457FB71A6B CRC64;

Query Match      87.5%; Score 42; DB 2; Length 144;
Best Local Similarity 87.5%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLPXP 12
```


RA Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum, and Eyeball;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK046962; BAC32927.1; -; mRNA.
DR EMBL; AK053696; BAC35479.1; -; mRNA.
DR MGI; MGI:2444096; Ppmlc.
SQ SEQUENCE 151 AA; 16372 MW; 868788C91FDD452E CRC64;

Query Match 87.5%; Score 42; DB 2; Length 151;
Best Local Similarity 87.5%; Pred. No. 2.le+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLXP 12
Db 120 PLPPLRP 127
|||||

RESULT 48
Q8PNK3_XANAC PRELIMINARY; PRT; 176 AA.
AC Q8PNK3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein XAC1056.
GN OrderedLocusNames=XAC1056;
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=22022145; PubMed=12024217; DOI=10.1038/417459a;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A.,
RA Almeida N.F. Jr., Alves L.M.C., do Amaral A.M., Bertolini M.C.,
RA Camargo L.E.A., Camarotte G., Cannavan F., Cardozo J., Chambergo F.,
RA Ciapina L.P., Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R.,
RA El-Dorrry H., Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Gruber A.,
RA Ferro M.I.T., Formighieri E.F., Franco M.C., Greggio C.C.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463(2002).

DR EMBL; AE011735; AAM35936.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 176 AA; 18964 MW; 14242128E08E65FF CRC64;

Query Match 87.5%; Score 42; DB 2; Length 176;
Best Local Similarity 87.5%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PLPPLXP 12
Db 50 PLPPLPKP 57
|||||

RESULT 49
Q9D7T0_MOUSE PRELIMINARY; PRT; 188 AA.
ID Q9D7T0_MOUSE PRELIMINARY;
AC Q9D7T0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Mus musculus adult male stomach cDNA, RIKEN full-length enriched
DE library, clone:2210411K11 product:hypothetical protein, full insert
DE sequence.
GN Name=2210411K11Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Stomach;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Stomach;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-Oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Stomach;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Stomach;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,

RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes."
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Stomach;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer."
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Stomach;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK008900; BAB25960.1; -; mRNA.
DR Ensembl; ENSMUSG00000030431; Mus musculus.
DR MGI; MGI:1922935; 2210411K11Rik.
KW Hypothetical protein.
SQ SEQUENCE 188 AA; 19274 MW; 4B6F8DAD2EDCB327 CRC64;

Query Match 87.5%; Score 42; DB 2; Length 188;
Best Local Similarity 77.8%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPLPPLXP 12
||| |||||
Db 147 RPSFPLPAP 155

RESULT 50
O01621 CAEEL
ID O01621_CAEEL PRELIMINARY; PRT; 198 AA.
AC O01621;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein W08A12.3.
GN ORFNames=W08A12.3;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL; U97010; AAB52319.1; -; Genomic_DNA.
DR PIR; E89008; E89008.
DR Ensembl; W08A12.3; Caenorhabditis elegans.
DR WormBase; WBGene0021080; W08A12.3.

DR WormPep; W08A12.3; CE14700.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 198 AA; 21928 MW; 7CFD2B55D123FCBF CRC64;
Query Match 87.5%; Score 42; DB 2; Length 198;
Best Local Similarity 87.5%; Pred. No. 2.8e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 5 PLPPLXP 12
||| |||||
Db 98 PLPPLPAP 105

Search completed: April 6, 2006, 09:38:32
Job time : 180.895 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.
OM protein - protein search, using sw model
Run on: April 6, 2006, 09:22:55 ; Search time 134.211 Seconds
(without alignments)
39.286 Million cell updates/sec

Title: US-10-632-388-312
Perfect score: 67
Sequence: 1 PPPYPPPPPIPX 12
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues 2443163

Total number of hits satisfying chosen parameters:
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score		Query Match	Length	DB ID	Description
	%					
1	67	100.0	10	2	AAR93314	Aar93314 ABL prote
2	67	100.0	10	6	ADA08234	Ada08234 Human Abl
3	67	100.0	12	2	AAR93369	Aar93369 ABL prote
4	67	100.0	12	2	AAR93371	Aar93371 ABL prote
5	67	100.0	12	3	AAB17256	Aab17256 SH3 antag
6	67	100.0	12	5	ABB73249	Abb73249 Src homol
7	67	100.0	12	7	ADJ73403	Adj73403 SH3 antag
8	67	100.0	12	8	ADJ53037	Adj53037 CH1 delet
9	67	100.0	12	8	ADJ51998	Adj51998 CH1 delet
10	66	98.5	12	2	AAR93373	Aar93373 ABL prote
11	66	98.5	12	2	AAR93372	Aar93372 ABL prote
12	66	98.5	12	2	AAR93370	Aar93370 ABL prote
13	66	98.5	12	3	AAB17257	Aab17257 SH3 antag
14	66	98.5	12	5	ABB73250	Abb73250 Src homol
15	66	98.5	12	7	ADJ73404	Adj73404 SH3 antag
16	66	98.5	12	8	ADJ53038	Adj53038 CH1 delet
17	66	98.5	12	8	ADJ51999	Adj51999 CH1 delet
18	60	89.6	10	2	AAR93315	Aar93315 ABL prote
19	59	88.1	10	2	AAR93318	Aar93318 ABL prote
20	59	88.1	10	2	AAR93319	Aar93319 ABL prote
21	59	88.1	11	2	AAR93374	Aar93374 ABL prote
22	59	88.1	12	2	AAR93376	Aar93376 ABL prote
23	58	86.6	10	2	AAR93317	Aar93317 ABL prote
24	58	86.6	10	2	AAR93323	Aar93323 ABL prote

25	58	86.6	10	4	AAB98984	Aab98984 Yeast ORF
26	58	86.6	10	7	ADK70159	Adk70159 Human SH3
27	58	86.6	11	6	ABG75623	Abg75623 Human Abe
28	58	86.6	40	3	AAB12074	Aab12074 SH3 bindi
29	58	86.6	40	7	ADE84647	Ade84647 3BP2 SH3
30	58	86.6	467	8	ADQ67211	Adq67211 Novel hum
31	58	86.6	492	8	ADQ96504	Adq96504 T cell ac
32	58	86.6	536	8	ADQ66376	Adq66376 Novel hum
33	58	86.6	561	7	ADC71301	Adc71301 Human col
34	58	86.6	561	7	ADK70062	Adk70062 Mutant hu
35	58	86.6	561	7	ADK70058	Adk70058 Wild type
36	58	86.6	561	8	ADQ96506	Adq96506 T cell ac
37	57	85.1	50	4	ABG19592	Abg19592 Novel hum
38	57	85.1	94	4	ABB71558	Abb71558 Drosophil
39	57	85.1	98	4	AAO02495	Aao02495 Human pol
40	57	85.1	118	4	ABG16489	Abg16489 Novel hum
41	57	85.1	142	7	ABO76127	AbO76127 Pseudomon
42	57	85.1	182	4	AAB80893	Aab80893 Human ATF
43	57	85.1	216	4	AAB80892	Aab80892 Human ATF
44	57	85.1	216	4	ABB50151	Abb50151 Human tra
45	57	85.1	227	4	ABG19594	Abg19594 Novel hum
46	57	85.1	282	4	AAB80896	Aab80896 Human ATF
47	57	85.1	282	4	AAB95859	Aab95859 Human pro
48	57	85.1	282	5	ABB81169	Abb81169 Human act
49	57	85.1	282	7	ADE60923	Ade60923 Human pro
50	57	85.1	282	8	ADP12591	Adp12591 Protein e
51	57	85.1	282	8	ADQ20824	Adq20824 Human sof
52	57	85.1	282	8	ADR14371	Adr14371 Human NF-
53	57	85.1	282	8	ABM82139	Abm82139 Tumour-as
54	57	85.1	308	6	ABR41338	Abr41338 Human DIT
55	57	85.1	315	8	ADT57494	Adt57494 Plant pol
56	57	85.1	397	8	ADQ97684	Adq97684 Mouse can
57	57	85.1	417	4	ABG12837	Abg12837 Novel hum
58	57	85.1	557	8	ADY24937	Ady24937 Plant ful
59	57	85.1	783	2	AAW37151	Aaw37151 Mouse neu
60	57	85.1	787	2	AAW37152	Aaw37152 Mouse neu
61	57	85.1	801	8	ADQ97686	Adq97686 Mouse can
62	57	85.1	802	2	AAW37153	Aaw37153 Mouse neu
63	57	85.1	802	4	AAU09139	Aau09139 Mammalian
64	57	85.1	1134	5	AAE24341	Aae24341 Human lun
65	57	85.1	1134	7	ADF74123	Adf74123 Human nov
66	56	83.6	64	4	AAO00314	Aao00314 Human pol
67	56	83.6	104	4	AAM15554	Aam15554 Peptide #
68	56	83.6	104	4	ABB34560	Abb34560 Peptide #
69	56	83.6	104	4	AAM28043	Aam28043 Peptide #
70	56	83.6	104	4	ABB29387	Abb29387 Peptide #
71	56	83.6	104	4	ABB19968	Abb19968 Protein #
72	56	83.6	104	4	AAM67743	Aam67743 Human bon
73	56	83.6	104	4	AAM55347	Aam55347 Human bra
74	56	83.6	104	4	ABG49381	Abg49381 Human liv
75	56	83.6	104	4	AAO03303	Aao03303 Peptide #
76	56	83.6	104	5	ABG37302	Abg37302 Human pep
77	56	83.6	113	5	ADK34538	Adk34538 Novel hum
78	56	83.6	156	4	ABG14149	Abg14149 Novel hum
79	56	83.6	280	8	ADY11773	Ady11773 Plant ful
80	56	83.6	293	8	ADM97712	Adm97712 Human MNA
81	56	83.6	362	3	AAB42278	Aab42278 Human ORF
82	56	83.6	368	8	ADM97711	Adm97711 Human MNA
83	56	83.6	386	6	ABR47582	Abr47582 Breast ca
84	56	83.6	393	2	AAW37149	Aaw37149 Murine En
85	56	83.6	393	4	AAU09148	Aau09148 Ena/VASP-
86	56	83.6	393	7	ADE62912	Ade62912 Rat Prote
87	56	83.6	393	7	ADD48812	Add48812 Rat Prote
88	56	83.6	393	7	ADG46666	Adg46666 Mouse ena
89	56	83.6	393	8	ADJ14191	Adj14191 Mouse ena
90	56	83.6	418	6	ABR47581	Abr47581 Breast ca
91	56	83.6	418	7	ADE62914	Ade62914 Human Pro
92	56	83.6	418	7	ADD48814	Add48814 Human Pro
93	56	83.6	418	8	ADP55192	Adp55192 Human PRO
94	56	83.6	418	8	ADP24828	Adp24828 PRO polyp
95	56	83.6	418	9	ADY17279	Ady17279 PRO polyp
96	56	83.6	418	9	ADY70308	Ady70308 Human bet
97	56	83.6	456	4	AAU17420	Aau17420 Novel sig

98	56	83.6	456	7	ADB94128	Human nov	171	55	82.1	267	8	ADJ63977	Adj63977 Human Fas
99	56	83.6	487	4	ABG21207	Novel hum	172	55	82.1	267	8	ADL71817	Adl71817 Human Fas
100	56	83.6	536	8	ADM97710	Human MNA	173	55	82.1	267	8	ADN07588	Adn07588 Human Fas
101	56	83.6	582	4	ABB59603	Drosophil	174	55	82.1	268	2	AAW48953	Aaw48953 Non-cleav
102	56	83.6	745	5	AAE19569	Human cyt	175	55	82.1	271	2	AAy28596	Aay28596 Fas ligand
103	56	83.6	777	5	ABP41839	Human ova	176	55	82.1	277	2	AAy04372	Aay04372 Human Fas
104	56	83.6	880	9	ADX83085	Human PEL	177	55	82.1	277	2	AAy28595	Aay28595 Fas ligand
105	56	83.6	905	2	AAW31186	Human p16	178	55	82.1	278	6	ABR82203	Abr82203 Chimeric
106	56	83.6	945	4	ABB60000	Drosophil	179	55	82.1	280	8	ADJ36211	Adj36211 Self-coal
107	56	83.6	983	8	ADM87342	Human pro	180	55	82.1	281	2	AAR77281	Aar77281 Human Fas
108	56	83.6	1021	7	ADJ68507	Human hea	181	55	82.1	281	2	AAR79097	Aar79097 Human Fas
109	56	83.6	1130	7	ADC99065	Human KPP	182	55	82.1	281	2	AAR88356	Aar88356 Human Fas
110	56	83.6	1130	8	ADM97695	Human MNA	183	55	82.1	281	2	AAR98104	Aar98104 Human Fas
111	56	83.6	1130	8	ABM81812	Tumour-as	184	55	82.1	281	2	AAW27143	Aaw27143 Human Fas
112	56	83.6	1130	9	ADX83072	Proline,	185	55	82.1	281	2	AAW75959	Aaw75959 Human Fas
113	56	83.6	1135	2	AAW31185	Human p16	186	55	82.1	281	2	AAW49105	Aaw49105 Fas Ligand
114	56	83.6	1251	7	ADE09332	Novel pro	187	55	82.1	281	2	AAW98071	Aaw98071 Human Fas
115	56	83.6	1251	8	ADM87788	Human EST	188	55	82.1	281	2	AAy04373	Aay04373 Human Fas
116	56	83.6	1493	4	ABB69228	Drosophil	189	55	82.1	281	2	AAW95041	Aaw95041 Human Fas
117	56	83.6	1953	6	ABR53597	Protein s	190	55	82.1	281	2	AAy28594	Aay28594 Wild type
118	56	83.6	1953	7	ADK63514	Disease t	191	55	82.1	281	3	AAy87580	Aay87580 Human Fas
119	56	83.6	1953	8	ADN19292	Bacterial	192	55	82.1	281	3	AAy87577	Aay87577 Human Fas
120	56	83.6	1958	2	AAR60620	Protein f	193	55	82.1	281	3	AAy87569	Aay87569 Human Fas
121	55	82.1	10	4	AAB86131	Proline-r	194	55	82.1	281	3	AAy87575	Aay87575 Human Fas
122	55	82.1	10	4	AAB86149	Proline-r	195	55	82.1	281	3	AAy87576	Aay87576 Human Fas
123	55	82.1	10	4	AAB70933	Polyoma v	196	55	82.1	281	3	AAy87579	Aay87579 Human Fas
124	55	82.1	13	4	AAB70928	Polyoma v	197	55	82.1	281	3	AAy87574	Aay87574 Human Fas
125	55	82.1	14	2	AAW38060	PPPY mot	198	55	82.1	281	3	AAy87582	Aay87582 Human Fas
126	55	82.1	14	7	ADB49250	Biotinyla	199	55	82.1	281	3	AAy87578	Aay87578 Human Fas
127	55	82.1	18	2	AAW38923	Peptide r	200	55	82.1	281	3	AAy87581	Aay87581 Human Fas
128	55	82.1	18	2	AAW39010	Peptide r	201	55	82.1	281	3	AAy87581	Aay87581 Human Fas
129	55	82.1	20	2	AAW38961	Peptide r	202	55	82.1	281	4	AAU04556	Aau04556 Human Fas
130	55	82.1	28	2	AAy41643	Mammalian	203	55	82.1	281	5	ABB80000	Abb80000 Human Fas
131	55	82.1	29	6	AAE30902	Novel pro	204	55	82.1	281	5	ABG96462	Abg96462 Human Fas
132	55	82.1	46	4	ABG12808	Novel hum	205	55	82.1	281	5	ABB05000	Abb05000 Human Fas
133	55	82.1	59	3	AAG02396	Human sec	206	55	82.1	281	5	ABB81652	Abb81652 Human Fas
134	55	82.1	85	4	AAO02036	Human pol	207	55	82.1	281	5	AAE29153	Aae29153 Wild-type
135	55	82.1	85	4	AAO04573	Human pol	208	55	82.1	281	6	ABR42309	Abr42309 Human Fas
136	55	82.1	93	8	ADR94152	Novel S.	209	55	82.1	281	6	ABP60549	Abp60549 Human tum
137	55	82.1	93	9	AEA58022	Streptoco	210	55	82.1	281	7	ADC35194	Adc35194 Human TNF
138	55	82.1	102	5	ABP01651	Human ORF	211	55	82.1	281	7	ADF44995	Adf44995 Human Fas
139	55	82.1	121	8	ADR10352	Human pro	212	55	82.1	281	7	ABW02272	Abw02272 Human Fas
140	55	82.1	127	5	AAE29154	Human non	213	55	82.1	281	7	ADG14422	Adg14422 Human Fas
141	55	82.1	137	3	AAG18806	Arabidops	214	55	82.1	281	7	ADJ74024	Adj74024 Human Fas
142	55	82.1	138	3	AAG18805	Arabidops	215	55	82.1	281	7	ADJ92614	Adj92614 Human Fas
143	55	82.1	139	4	ABB41888	Peptide #	216	55	82.1	281	7	ADL17694	Adl17694 Human Fas
144	55	82.1	139	4	AAW35687	Peptide #	217	55	82.1	281	8	ADP90414	Adp90414 Human Fas
145	55	82.1	139	4	AAW75582	Human bon	218	55	82.1	281	8	ADJ56767	Adj56767 Human Fas
146	55	82.1	139	4	AAW62764	Human bra	219	55	82.1	281	8	ADL27664	Adl27664 Human Fas
147	55	82.1	139	4	ABG57324	Human liv	220	55	82.1	281	8	ADL23577	Adl23577 Human Fas
148	55	82.1	139	5	ABG45112	Human pep	221	55	82.1	281	8	ADK19674	Adk19674 Chicken a
149	55	82.1	158	4	AAW65985	Human sec	222	55	82.1	281	8	ADK19673	Adk19673 Human Fas
150	55	82.1	158	4	AAW65986	Human sec	223	55	82.1	281	8	ADM53436	Adm53436 Human Fas
151	55	82.1	166	8	ADY05758	Plant ful	224	55	82.1	281	8	ADP12464	Adp12464 Protein e
152	55	82.1	166	8	ADX90309	Plant ful	225	55	82.1	281	8	ADO19786	Ado19786 Human PRO
153	55	82.1	172	7	ABM87038	Rice abio	226	55	82.1	281	8	ADO19800	Ado19800 Human PRO
154	55	82.1	173	7	ABO75054	Pseudomon	227	55	82.1	281	8	ADP26984	Adp26984 Human Fas
155	55	82.1	192	4	AAW48292	Human ZF4	228	55	82.1	281	8	ADS87994	Ads87994 Tumour tr
156	55	82.1	192	8	ADO62766	Transcrip	229	55	82.1	281	8	ADP56004	Adp56004 Human PRO
157	55	82.1	197	3	AAy83079	F-box pro	230	55	82.1	281	8	ADR44815	Adr44815 Fas ligand
158	55	82.1	197	5	AAO22465	Human F-b	231	55	82.1	281	8	ADR47163	Adr47163 Human Fas
159	55	82.1	197	9	ADY62389	Human F-b	232	55	82.1	281	8	ADT78394	Adt78394 Human Fas
160	55	82.1	234	3	AAB32786	Eucalyptu	233	55	82.1	281	8	ADU22865	Adu22865 Human apo
161	55	82.1	236	4	AAW39701	Human pol	234	55	82.1	281	9	ADY19512	Ady19512 PRO polyp
162	55	82.1	239	4	ABB70310	Drosophil	235	55	82.1	281	9	ADY16498	Ady16498 PRO polyp
163	55	82.1	248	9	ADW17139	Eucalyptu	236	55	82.1	281	9	ADZ14447	Adz14447 Human Fas
164	55	82.1	255	4	AAB86143	A. victor	237	55	82.1	288	8	ADK00625	Adk00625 HOMO prot
165	55	82.1	255	4	AAB70926	GFP-PLP c	238	55	82.1	303	8	ADI43195	Adi43195 Plant tra
166	55	82.1	258	2	AAy04371	Human Fas	239	55	82.1	303	8	ADO3266	Ado3266 Thalecres
167	55	82.1	258	7	ADG14424	Human Fas	240	55	82.1	310	6	ADA55489	Ada55489 Human pro
168	55	82.1	260	7	ADJ69950	Human hea	241	55	82.1	311	2	AAy17387	Aay17387 Human DNA
169	55	82.1	261	2	AAy28597	Fas ligand	242	55	82.1	323	7	ADM05940	Adm05940 Human pro
170	55	82.1	265	2	AAW48954	Non-cleav	243	55	82.1	326	3	AAG43482	Aag43482 Arabidops

244	55	82.1	326	3	AAG24964	Aag24964 Arabidops
245	55	82.1	328	7	ADD26784	Add26784 Human adi
246	55	82.1	329	8	ADJ36255	Adj36255 Self-coal
247	55	82.1	360	4	AAB93436	Aab93436 Human pro
248	55	82.1	360	8	ADN05477	Adn05477 Antipsori
249	55	82.1	360	8	ADN05382	Adn05382 Antipsori
250	55	82.1	360	9	ADY18819	Ady18819 PRO polyp
251	55	82.1	360	9	ADY79929	Ady79929 Amino aci
252	55	82.1	362	4	ABG16456	Abg16456 Novel hum
253	55	82.1	375	7	ADE09101	Ade09101 Novel hum
254	55	82.1	375	9	ADU40616	Adu40616 Novel hum
255	55	82.1	393	4	AAB70924	Aab70924 Polyoma v
256	55	82.1	419	8	ADR66870	Adr66870 Human pro
257	55	82.1	419	8	ADR65972	Adr65972 Human pro
258	55	82.1	426	4	ABB11048	Abb11048 Human equ
259	55	82.1	434	3	AAB42352	Aab42352 Human ORF
260	55	82.1	451	7	ADE62342	Ade62342 Rat Prote
261	55	82.1	454	7	ABM87734	Abm87734 Rice abio
262	55	82.1	456	6	ABP98836	Abp98836 Human str
263	55	82.1	467	7	ADM04921	Adm04921 Human pro
264	55	82.1	473	3	AAG24963	Aag24963 Arabidops
265	55	82.1	473	3	AAG43481	Aag43481 Arabidops
266	55	82.1	473	8	ADT55517	Adt55517 Plant pol
267	55	82.1	494	4	ABG29246	Abg29246 Novel hum
268	55	82.1	498	7	ABM86962	Abm86962 Rice abio
269	55	82.1	512	8	ADY13503	Ady13503 Plant ful
270	55	82.1	515	8	ADS12177	Ads12177 Human the
271	55	82.1	527	4	AAB70917	Aab70917 L. monocy
272	55	82.1	533	4	AAB95397	Aab95397 Human pro
273	55	82.1	541	2	AAW37148	Aaw37148 Mammalian
274	55	82.1	542	8	ADX79745	Adx79745 Plant ful
275	55	82.1	552	2	AAW90172	Aaw90172 Human hea
276	55	82.1	552	4	AAB86278	Aab86278 Human DCM
277	55	82.1	552	5	AAE25975	Aae25975 Human pro
278	55	82.1	552	7	ADH62560	Adh62560 Cardiac m
279	55	82.1	552	9	AEA21095	Aea21095 Novel hum
280	55	82.1	554	6	ADA48166	Ada48166 Rice prot
281	55	82.1	559	4	AAB20494	Aab20494 Human pab
282	55	82.1	559	4	AAB73399	Aab73399 Human WAV
283	55	82.1	559	4	AAG67345	Aag67345 Amino aci
284	55	82.1	559	4	AAM52317	Aam52317 Human Sca
285	55	82.1	559	7	ADD08985	Add08985 Human pab
286	55	82.1	559	7	ADJ68719	Adj68719 Human hea
287	55	82.1	559	7	ADJ70867	Adj70867 Human hea
288	55	82.1	559	8	ADG65216	Adg65216 Mouse Pab
289	55	82.1	559	8	ADG65208	Adg65208 Human Pab
290	55	82.1	559	8	ADR16249	Adr16249 Human Pab
291	55	82.1	559	8	ADS34608	Ads34608 POSH prot
292	55	82.1	568	6	ADA54994	Ada54994 Human pro
293	55	82.1	575	8	ADT07559	Adt07559 Human col
294	55	82.1	577	8	ADT07561	Adt07561 Human col
295	55	82.1	585	8	ADT07557	Adt07557 Human col
296	55	82.1	601	7	ADB64412	Adb64412 Human pro
297	55	82.1	619	4	ABB68682	Abb68682 Drosophil
298	55	82.1	629	4	AAY72168	Aay72168 Human RNA
299	55	82.1	634	7	ADM04327	Adm04327 Human pro
300	55	82.1	638	8	ADI82510	Adi82510 Human mod
301	55	82.1	656	9	ADV99949	Adv99949 Human PEM
302	55	82.1	659	9	ADV99948	Adv99948 Human PEM
303	55	82.1	659	9	ADV99945	Adv99945 Human PEM
304	55	82.1	675	6	AAO26721	Aao26721 SR protei
305	55	82.1	681	7	ADE59481	Ade59481 Rat Prote
306	55	82.1	686	5	AAU80373	Aau80373 Human cel
307	55	82.1	686	6	AAE34823	Aae34823 Protein #
308	55	82.1	686	8	ADO19347	Ado19347 Human PRO
309	55	82.1	686	9	ADX05467	Adx05467 Cyclin-de
310	55	82.1	686	9	ADY15012	Ady15012 PRO polyp
311	55	82.1	686	9	AEB34464	Aeb34464 SR protei
312	55	82.1	687	8	ADY07407	Ady07407 Plant ful
313	55	82.1	688	2	AAAY27053	Aay27053 Human pro
314	55	82.1	688	6	AAO26722	Aao26722 SR protei
315	55	82.1	688	8	ADJ96645	Adj96645 Human SRP
316	55	82.1	691	7	ADF76663	Adf76663 Novel hum

317	55	82.1	699	5	ABG32447	Abg32447 Human SR
318	55	82.1	699	7	ADB64660	Adb64660 Human pro
319	55	82.1	699	7	ADJ37689	Adj37689 Human kin
320	55	82.1	699	8	ADR31218	Adr31218 Human SRP
321	55	82.1	706	7	ADM05234	Adm05234 Human pro
322	55	82.1	715	8	ADX91718	Adx91718 Plant ful
323	55	82.1	737	9	ADV99974	Adv99974 Human PEM
324	55	82.1	737	9	ADV99972	Adv99972 Human PEM
325	55	82.1	749	7	ADB64468	Adb64468 Human pro
326	55	82.1	753	7	ADG31678	Adg31678 Human pro
327	55	82.1	798	5	AAU93051	Aau93051 Arabidops
328	55	82.1	798	7	ABO43116	Abo43116 A. thalia
329	55	82.1	798	7	ADD30228	Add30228 Plant yie
330	55	82.1	798	7	ADE37127	Ade37127 Plant yie
331	55	82.1	798	8	ADI41767	Adi41767 Plant tra
332	55	82.1	798	8	ADI61347	Adi61347 A. thalia
333	55	82.1	798	8	ADO02319	Ado02319 Thalecres
334	55	82.1	834	7	ADC31026	Adc31026 Human nov
335	55	82.1	844	8	ADR08920	Adr08920 Human pro
336	55	82.1	848	5	ABG66688	Abg66688 Human nov
337	55	82.1	849	6	ADA13338	Ada13338 Human int
338	55	82.1	854	3	AAG42329	Aag42329 Arabidops
339	55	82.1	854	7	ADD46674	Add46674 Human Pro
340	55	82.1	906	5	ABG66687	Abg66687 Human nov
341	55	82.1	907	6	ABG74229	Abg74229 Canine An
342	55	82.1	971	7	ADE56304	Ade56304 Human Pro
343	55	82.1	983	2	AAAY09513	Aay09513 Mouse JMY
344	55	82.1	986	7	ADD01164	Add01164 Human nuc
345	55	82.1	996	9	AEA20282	Aea20282 Novel hum
346	55	82.1	1001	6	ABU89812	Abu89812 Novel hum
347	55	82.1	1001	8	ADO09372	Ado09372 Novel hum
348	55	82.1	1005	2	AAW93955	Aaw93955 Human 53B
349	55	82.1	1005	6	ABU89810	Abu89810 Novel hum
350	55	82.1	1005	6	ABG71757	Abg71757 Human can
351	55	82.1	1005	6	ADA10593	Ada10593 Human can
352	55	82.1	1005	8	ADO09368	Ado09368 Novel hum
353	55	82.1	1005	9	ADX05602	Adx05602 Cyclin-de
354	55	82.1	1044	8	ADR09546	Adr09546 Human pro
355	55	82.1	1051	3	AAG42328	Aag42328 Arabidops
356	55	82.1	1051	8	ADT55608	Adt55608 Plant pol
357	55	82.1	1057	3	AAG42327	Aag42327 Arabidops
358	55	82.1	1068	6	ABU89811	Abu89811 Novel hum
359	55	82.1	1068	8	ADO09370	Ado09370 Novel hum
360	55	82.1	1070	4	ABG02155	Abg02155 Novel hum
361	55	82.1	1082	7	ADE14368	Ade14368 Human int
362	55	82.1	1088	8	ADH61284	Adh61284 INTSIG pr
363	55	82.1	1101	8	ADR66324	Adr66324 Human pro
364	55	82.1	1101	8	ADR66666	Adr66666 Human pro
365	55	82.1	1112	8	ADR66113	Adr66113 Human pro
366	55	82.1	1112	8	ADR666455	Adr666455 Human pro
367	55	82.1	1112	9	ADY18523	Ady18523 PRO polyp
368	55	82.1	1128	9	ADN12357	Adn12357 ASPP2 pol
369	55	82.1	1128	9	AEA32205	Aea32205 Human ASP
370	55	82.1	1128	9	AEB16419	Aeb16419 ASPP2. 8/
371	55	82.1	1141	6	ABR39812	Abx39812 Human SCA
372	55	82.1	1143	7	ADC31787	Adc31787 Human nov
373	55	82.1	1152	9	ADY61687	Ady61687 Full leng
374	55	82.1	1171	5	AAE16987	Aae16987 Mouse DRF
375	55	82.1	1193	8	ADP04178	Adp04178 Human col
376	55	82.1	1196	6	ABR39811	Abx39811 Human SCA
377	55	82.1	1203	8	ADT07558	Adt07558 Human col
378	55	82.1	1205	8	ADT07560	Adt07560 Human col
379	55	82.1	1213	8	ADT07556	Adt07556 Human col
380	55	82.1	1216	8	ADX97588	Adx97588 Pancreati
381	55	82.1	1248	2	AAAY13464	Aay13464 Human dia
382	55	82.1	1250	6	AAO26231	Aao26231 MDDT rela
383	55	82.1	1250	8	ADM87156	Adm87156 Human pro
384	55	82.1	1255	2	AAW52249	Aaw52249 Mouse Rho
385	55	82.1	1255	2	AAW76733	Aaw76733 Mouse mDi
386	55	82.1	1315	2	AAW76734	Aaw76734 Human mDi
387	55	82.1	1331	4	ABG28241	Abg28241 Novel hum
388	55	82.1	1375	8	ADS43646	Ads43646 Bacterial
389	55	82.1	1410	8	ABM81569	Abm81569 Tumour-as

390	55	82.1	1467	5	AAU78283	Aau78283 Human apo	463	52	77.6	41	8	ADU09353	Adu09353 Synthetic
391	55	82.1	1527	4	AAU01184	Aau01184 Rat gluc	464	52	77.6	44	4	ABG16321	Abg16321 Novel hum
392	55	82.1	1527	8	ADF45523	Adf45523 Rat KIAA0	465	52	77.6	46	4	ABG17583	Abg17583 Novel hum
393	55	82.1	1592	8	ADL35711	Adl35711 Human ang	466	52	77.6	47	5	ADR41538	Adr41538 Human CD-
394	55	82.1	1592	8	ABM81997	Abm81997 Tumour-as	467	52	77.6	53	6	ABB98725	Abb98725 Human PR1
395	55	82.1	1616	8	ADP22958	Adp22958 PRO poly	468	52	77.6	53	6	ABB98733	Abb98733 Murine PR
396	55	82.1	1652	7	ADE71273	Ade71273 Novel hum	469	52	77.6	54	4	ABG05172	Abg05172 Novel hum
397	55	82.1	1682	8	ADR14131	Adr14131 Human NF-	470	52	77.6	69	8	ADR94153	Adr94153 Novel S.
398	55	82.1	1765	7	ADB80341	Adb80341 Human MDD	471	52	77.6	69	9	AEA58023	Aea58023 Streptoco
399	55	82.1	1822	5	ABP43899	Abp43899 Nuclear p	472	52	77.6	78	6	ABB98736	Abb98736 Murine PR
400	55	82.1	1822	7	ADE59297	Ade59297 Human Pro	473	52	77.6	78	6	ABB98728	Abb98728 Human PR1
401	55	82.1	1822	7	ADE59293	Ade59293 Human Pro	474	52	77.6	81	5	ABP34010	Abp34010 Human ORF
402	55	82.1	2062	7	ADC31187	Adc31187 Human nov	475	52	77.6	88	4	AAO02124	Aao02124 Human pol
403	55	82.1	2068	8	ADO62867	Ado62867 Transcrip	476	52	77.6	88	6	ABB98739	Abb98739 Murine PR
404	55	82.1	2150	9	ADZ09836	Adz09836 Human bre	477	52	77.6	88	6	ABB98731	Abb98731 Human PR1
405	55	82.1	2783	2	AAR23963	Aar23963 AFP-1 (Al	478	52	77.6	103	3	AAB33096	Aab33096 Pinus rad
406	55	82.1	2783	2	AAR23962	Aar23962 AFP-1. 3/	479	52	77.6	104	5	ABB78538	Abb78538 Ser-Pro-P
407	55	82.1	2783	4	AAB82946	Aab82946 Human and	480	52	77.6	104	8	ADU09325	Adu09325 [SPPPP]n
408	55	82.1	2789	7	ADJ70422	Adj70422 Human hea	481	52	77.6	105	4	AAO00771	Aao00771 Human pol
409	55	82.1	2992	8	ADP30190	Adp30190 Human sec	482	52	77.6	113	3	AAG11750	Aag11750 Arabidops
410	55	82.1	3065	8	ADP30259	Adp30259 Human sec	483	52	77.6	113	6	ABB98732	Abb98732 Human PR1
411	55	82.1	3394	7	ADJ68723	Adj68723 Human hea	484	52	77.6	113	6	ABB98740	Abb98740 Murine PR
412	55	82.1	3572	5	ABG95659	Abg95659 Human nuc	485	52	77.6	114	7	ADC33200	Adc33200 Human nov
413	55	82.1	4873	6	ABO14747	Abol14747 Novel hum	486	52	77.6	118	6	ABB98730	Abb98730 Human PR1
414	54	80.6	171	7	ADF08027	Adf08027 Bacterial	487	52	77.6	118	6	ABB98738	Abb98738 Murine PR
415	54	80.6	242	8	ADY22812	Ady22812 Plant ful	488	52	77.6	121	2	AAR29164	Aar29164 PRP3. 3/2
416	54	80.6	378	8	ADX72267	Adx72267 Plant ful	489	52	77.6	127	3	AAG06240	Aag06240 Arabidops
417	54	80.6	406	6	ABP58350	Abp58350 Human cel	490	52	77.6	127	3	AAG40268	Aag40268 Arabidops
418	54	80.6	485	6	ABP58349	Abp58349 Human cel	491	52	77.6	128	3	AAG40267	Aag40267 Arabidops
419	54	80.6	485	7	ADC31623	Adc31623 Human nov	492	52	77.6	128	3	AAG06239	Aag06239 Arabidops
420	54	80.6	485	7	ADM05364	Adm05364 Human pro	493	52	77.6	128	3	AAG07010	Aag07010 Arabidops
421	54	80.6	591	4	ABG03618	Abg03618 Novel hum	494	52	77.6	131	6	ABJ25281	Abj25281 Mouse BAC
422	54	80.6	722	8	ADQ65219	Adq65219 Novel hum	495	52	77.6	132	4	AAO05753	Aao05753 Human pol
423	54	80.6	825	7	ADJ70479	Adj70479 Human hea	496	52	77.6	133	4	ABG09692	Abg09692 Novel hum
424	54	80.6	900	3	AAB42321	Aab42321 Human ORF	497	52	77.6	137	4	ABG12622	Abg12622 Novel hum
425	54	80.6	1005	8	ADQ65183	Adq65183 Novel hum	498	52	77.6	138	3	AAG11749	Aag11749 Arabidops
426	54	80.6	1091	4	ABB58383	Abb58383 Drosophil	499	52	77.6	138	4	ABG22949	Abg22949 Novel hum
427	54	80.6	1217	4	ABG09876	Abg09876 Novel hum	500	52	77.6	138	7	ADC33086	Adc33086 Human nov
428	54	80.6	1638	7	ADC31236	Adc31236 Human nov	501	52	77.6	140	8	ADX73862	Adx73862 Plant ful
429	54	80.6	1822	8	ADN21439	Adn21439 Bacterial	502	52	77.6	144	3	AAG40266	Aag40266 Arabidops
430	54	80.6	2397	8	ADU02443	Adu02443 Novel hum	503	52	77.6	144	3	AAG07009	Aag07009 Arabidops
431	53	79.1	29	4	AAO04677	Aao04677 Human pol	504	52	77.6	144	3	AAG06238	Aag06238 Arabidops
432	53	79.1	56	4	ABG12672	Abg12672 Novel hum	505	52	77.6	151	2	AAR29165	Aar29165 PRP3 (fro
433	53	79.1	100	4	ABG25413	Abg25413 Novel hum	506	52	77.6	152	7	ADB75611	Adb75611 Prostate
434	53	79.1	149	4	AAO05346	Aao05346 Human pol	507	52	77.6	153	1	AAP90380	Aap90380 Sequence
435	53	79.1	239	4	ABB63164	Abb63164 Drosophil	508	52	77.6	153	4	AAB67456	Aab67456 Amino aci
436	53	79.1	245	7	ADJ11494	Adj11494 Rice prot	509	52	77.6	153	6	ABB98723	Abb98723 Human PR1
437	53	79.1	245	7	ADJ11826	Adj11826 Rice prot	510	52	77.6	153	6	ABB98724	Abb98724 Murine PR
438	53	79.1	263	8	ADT57062	Adt57062 Plant pol	511	52	77.6	156	3	AAG07008	Aag07008 Arabidops
439	53	79.1	271	9	ABM96440	Abm96440 M. xanthu	512	52	77.6	158	4	AAB67457	Aab67457 Amino aci
440	53	79.1	363	4	AAM25490	Aam25490 Human pro	513	52	77.6	161	6	ABU19480	Abu19480 Protein e
441	53	79.1	400	7	ADC29743	Adc29743 B. juncea	514	52	77.6	167	3	AAG11748	Aag11748 Arabidops
442	53	79.1	493	4	ABG12673	Abg12673 Novel hum	515	52	77.6	175	8	ADN42048	Adn42048 Amino aci
443	53	79.1	511	8	ADR94269	Adr94269 Novel S.	516	52	77.6	178	8	ADN42050	Adn42050 Amino aci
444	53	79.1	511	9	AEA58139	Aea58139 Streptoco	517	52	77.6	184	8	ABO60210	Abo60210 Human gen
445	53	79.1	1258	3	AAY80120	Aay80120 Human Shi	518	52	77.6	189	5	ABG76914	Abg76914 Rice WRKY
446	53	79.1	1258	4	AAB98987	Aab98987 Human typ	519	52	77.6	198	4	ABG19714	Abg19714 Novel hum
447	53	79.1	1258	8	ADJ66611	Adj66611 Inositol	520	52	77.6	203	7	ADF08002	Adf08002 Bacterial
448	53	79.1	1258	8	ABM81289	Abm81289 Tumour-as	521	52	77.6	205	6	ABU20105	Abu20105 Protein e
449	53	79.1	1302	8	ADE28331	Ade28331 Human KPP	522	52	77.6	206	4	AAM42138	Aam42138 Human pol
450	53	79.1	1316	4	ABG22997	Abg22997 Novel hum	523	52	77.6	211	7	ADD30995	Add30995 Plant yie
451	53	79.1	1321	4	ABG25416	Abg25416 Novel hum	524	52	77.6	211	8	ADI43759	Adi43759 Plant tra
452	52	77.6	10	2	AAR93322	Aar93322 ABL prote	525	52	77.6	211	8	ADO01627	Ado01627 Thalecres
453	52	77.6	10	5	ABB78428	Abb78428 Hydroxypr	526	52	77.6	220	2	AAY21431	Aay21431 Human hig
454	52	77.6	10	8	ADU09142	Adu09142 Ser-Hyp4	527	52	77.6	228	4	AAU05775	Aau05775 Rice inve
455	52	77.6	19	1	AAP81947	Aap81947 19 amino	528	52	77.6	230	8	ADU09357	Adu09357 Synthetic
456	52	77.6	20	2	AAW38960	Aaw38960 Peptide r	529	52	77.6	233	3	AAB07859	Aab07859 Amino aci
457	52	77.6	20	8	ADU79115	Adu79115 Human pep	530	52	77.6	245	3	AAB43357	Aab43357 Human ORF
458	52	77.6	29	4	AAO02160	Aao02160 Human pol	531	52	77.6	245	8	ADF55743	Adp55743 Human PRO
459	52	77.6	29	4	ABG08613	Abg08613 Novel hum	532	52	77.6	257	2	AAY25630	Aay25630 Phleum sp
460	52	77.6	30	4	AAO02051	Aao02051 Human pol	533	52	77.6	257	7	ADC34880	Adc34880 Timothy g
461	52	77.6	33	3	AAY54915	Aay54915 RGD-contr	534	52	77.6	264	4	AAB64417	Aab64417 Amino aci
462	52	77.6	41	2	AAY01285	Aay01285 Peptide e	535	52	77.6	264	6	ADA54869	Ada54869 Human pro

536 52 77.6 264 7 ADM04325 Adm04325 Human pro
537 52 77.6 270 4 AAU16480 Aau16480 Human nov
538 52 77.6 270 6 ABU5549 Abu5549 Human nov
539 52 77.6 270 8 ADS11801 Adsl1801 Human the
540 52 77.6 271 9 ADY18585 Adv18585 PRO polyp
541 52 77.6 272 6 ABO14766 Abo14766 Novel hum
542 52 77.6 275 9 ADW17101 Adw17101 Eucalyptu
543 52 77.6 285 2 AAY25633 Aay25633 Phleum sp
544 52 77.6 285 7 ADC34883 Adc34883 Timothy g
545 52 77.6 293 3 AAB43674 Aab43674 Human can
546 52 77.6 297 8 ADR08753 Adr08753 Human pro
547 52 77.6 301 4 ABB68068 Abb68068 Drosophil
548 52 77.6 316 5 AAE222089 Aae222089 Human nov
549 52 77.6 319 8 ABO59236 Abo59236 Human gen
550 52 77.6 323 2 AAW00923 Aaw00923 HaSNPV po
551 52 77.6 325 7 ABM85700 Abm85700 Mouse pro
552 52 77.6 331 4 ABB70820 Abb70820 Drosophil
553 52 77.6 352 7 ADM04873 Adm04873 Human pro
554 52 77.6 353 5 AAE222090 Aae222090 Human nov
555 52 77.6 354 5 ABP41838 Abp41838 Human ova
556 52 77.6 356 4 AAB70701 Aab70701 Human TAA
557 52 77.6 357 5 AAU85555 Aau85555 Clone #50
558 52 77.6 357 5 AAU69431 Aau69431 Lung smal
559 52 77.6 357 6 ABU69527 Abu69527 Human lun
560 52 77.6 357 6 ABU66430 Abu66430 Lung canc
561 52 77.6 357 7 ADH47336 Adh47336 Human lun
562 52 77.6 357 8 ADJ21255 Adj21255 Human lun
563 52 77.6 360 8 ADY09822 Ady09822 Plant ful
564 52 77.6 390 3 AAB58287 Aab58287 Lung canc
565 52 77.6 391 7 ABM87129 Abm87129 Rice abio
566 52 77.6 402 9 ADX83401 Adx83401 Human TEG
567 52 77.6 403 8 ADX74016 Adx74016 Plant ful
568 52 77.6 418 4 AAB20496 Aab20496 Human Pab
569 52 77.6 418 6 ABU11825 Abu11825 Human MDD
570 52 77.6 418 8 ADG65212 Adg65212 Human Pab
571 52 77.6 418 8 ADG65220 Adg65220 Human Pab
572 52 77.6 421 6 AAO16115 Aao16115 Human can
573 52 77.6 428 3 AAG29606 Aag29606 Arabidops
574 52 77.6 431 3 AAG32324 Aag32324 Arabidops
575 52 77.6 431 3 AAG09613 Aag09613 Arabidops
576 52 77.6 431 8 ADI43535 Adi43535 Plant tra
577 52 77.6 431 9 AEA26765 Aea26765 Stress to
578 52 77.6 435 3 AAG09612 Aag09612 Arabidops
579 52 77.6 439 8 ADY25088 Ady25088 Plant ful
580 52 77.6 440 3 AAG49463 Aag49463 Arabidops
581 52 77.6 443 3 AAG49462 Aag49462 Arabidops
582 52 77.6 451 3 AAG35712 Aag35712 Arabidops
583 52 77.6 451 8 ADT56622 Adt56622 Plant pol
584 52 77.6 455 5 ABP73562 Abp73562 Candida a
585 52 77.6 470 4 ABG21932 Abg21932 Novel hum
586 52 77.6 480 4 ABG12623 Abg12623 Novel hum
587 52 77.6 483 7 ADU02754 Adu02754 Novel hum
588 52 77.6 487 7 ADJ91997 Adj91997 Human her
589 52 77.6 488 3 AAG22819 Aag22819 Arabidops
590 52 77.6 504 5 ABP41965 Abp41965 Human ova
591 52 77.6 520 4 AAB49336 Aab49336 Murine ova
592 52 77.6 520 4 AAG67350 Aag67350 Amino aci
593 52 77.6 520 4 AAM52318 Aam52318 Murine WA
594 52 77.6 521 7 ABM85832 Abm85832 Mouse pro
595 52 77.6 523 7 ADC87177 Adc87177 Human GPC
596 52 77.6 530 6 ADA55193 Ada55193 Human pro
597 52 77.6 551 8 ADQ65217 Adq65217 Novel hum
598 52 77.6 560 4 AAB95239 Aab95239 Human pro
599 52 77.6 594 4 ABB61362 Abb61362 Drosophil
600 52 77.6 599 7 ADM03784 Adm03784 Human pro
601 52 77.6 603 5 AAE222088 Aae222088 Human nov
602 52 77.6 603 8 ADJ65163 Adj65163 Potassium
603 52 77.6 620 9 ADY54945 Ady54945 Chronic v
604 52 77.6 622 8 ADJ65161 Adj65161 Potassium
605 52 77.6 649 7 ADC87385 Adc87385 Human GPC
606 52 77.6 649 7 ADD46407 Add46407 Human pro
607 52 77.6 651 9 ADY17208 Ady17208 PRO polyp
608 52 77.6 656 5 AAE19183 Aae19183 Human pro

609 52 77.6 666 2 AAW72911 Aaw72911 Mycobacte
610 52 77.6 666 2 AAY21928 Aay21928 Amino aci
611 52 77.6 666 4 AAB19846 Aab19846 Mycobacte
612 52 77.6 666 5 ABU05989 Abu05989 M. tuberc
613 52 77.6 666 6 ABU37020 Abu37020 Protein e
614 52 77.6 668 5 ABB93879 Abb93879 Herbicida
615 52 77.6 668 9 ADY91562 Ady91562 Thale cre
616 52 77.6 681 4 ABG07912 Abg07912 Novel hum
617 52 77.6 708 7 ADC19791 Adc19791 Human can
618 52 77.6 711 8 ADJ71971 Adj71971 Human PMM
619 52 77.6 721 8 ADY22716 Ady22716 Plant ful
620 52 77.6 730 8 ADO62718 Ado62718 Transcrip
621 52 77.6 781 8 ADO62717 Ado62717 Transcrip
622 52 77.6 800 8 ABO58564 Abo58564 Human gen
623 52 77.6 809 5 ABU03210 Abu03210 Human pho
624 52 77.6 809 8 ADI15651 Adi15651 Human pho
625 52 77.6 809 8 ADJ25986 Adj25986 Human pho
626 52 77.6 809 8 ADN97887 Adn97887 Human pho
627 52 77.6 809 8 ADO50282 Ado50282 Human pho
628 52 77.6 809 8 ADO34034 Ado34034 Human pho
629 52 77.6 809 8 ADQ78350 Adq78350 Human pho
630 52 77.6 809 9 AEB85186 Aeb85186 Human pho
631 52 77.6 828 4 ABB58368 Abb58368 Drosophil
632 52 77.6 847 8 ADR58600 Adr58600 Human int
633 52 77.6 863 4 AAB86463 Aab86463 Murine HC
634 52 77.6 885 7 ADE08478 Ade08478 Novel pro
635 52 77.6 913 9 AEA19955 Aea19955 Novel hum
636 52 77.6 1013 8 ADP24491 Adp24491 PRO polyp
637 52 77.6 1017 4 AAM40352 Aam40352 Human pol
638 52 77.6 1039 7 ADJ70453 Adj70453 Human hea
639 52 77.6 1039 7 ABM90450 Abm90450 Rice abio
640 52 77.6 1054 5 ABB79311 Abb79311 Human ova
641 52 77.6 1066 3 AAY84877 Aay84877 Amino aci
642 52 77.6 1076 5 ABP69070 Abp69070 Human pol
643 52 77.6 1098 7 ADE60625 Ade60625 Rat Prote
644 52 77.6 1100 7 ADE54778 Ade54778 Human Pro
645 52 77.6 1100 7 ADE54782 Ade54782 Human Pro
646 52 77.6 1100 8 ADN04034 Adn04034 Antipsori
647 52 77.6 1126 6 ABU70708 Abu70708 Human adi
648 52 77.6 1126 8 ABM80766 Abm80766 Tumour-as
649 52 77.6 1126 8 ABM80767 Abm80767 Tumour-as
650 52 77.6 1132 7 ADD71144 Add71144 Human int
651 52 77.6 1132 7 ADE60627 Ade60627 Human Pro
652 52 77.6 1132 7 ADE60623 Ade60623 Human Pro
653 52 77.6 1132 8 ABM80765 Abm80765 Tumour-as
654 52 77.6 1139 5 ABG97360 Abg97360 Human CGD
655 52 77.6 1161 5 ABG97354 Abg97354 Human CGD
656 52 77.6 1179 8 ADS10655 Adsl0655 Human the
657 52 77.6 1200 6 ADA55114 Ada55114 Human pro
658 52 77.6 1222 3 AAG36453 Aag36453 Arabidops
659 52 77.6 1239 4 ABG09877 Abg09877 Novel hum
660 52 77.6 1257 3 AAG36452 Aag36452 Arabidops
661 52 77.6 1275 3 AAG36451 Aag36451 Arabidops
662 52 77.6 1312 5 ABP69782 Abp69782 Human pol
663 52 77.6 1337 8 ABM82145 Abm82145 Tumour-as
664 52 77.6 1343 7 ADJ69279 Adj69279 Human hea
665 52 77.6 1362 8 ADJ66579 Adj66579 Bromodoma
666 52 77.6 1362 9 ADV85499 Adv85499 Human bro
667 52 77.6 1362 9 ADX07379 Adx07379 Cyclin-de
668 52 77.6 1386 7 ADE59291 Ade59291 Rat Prote
669 52 77.6 1386 7 ADE59295 Ade59295 Rat Prote
670 52 77.6 1390 5 ABU65152 Abu65152 Human NOV
671 52 77.6 1390 8 ADN61955 Adn61955 Human nov
672 52 77.6 1400 9 ADV85501 Adv85501 Mouse bro
673 52 77.6 1542 8 ADL83092 Adl83092 Human PRO
674 52 77.6 1542 8 ADN04196 Adn04196 Antipsori
675 52 77.6 1542 8 ADP56292 Adp56292 Human PRO
676 52 77.6 1542 9 ADY15650 Ady15650 PRO polyp
677 52 77.6 1555 4 ABB58847 Abb58847 Drosophil
678 52 77.6 1596 7 AAE37929 Aae37929 Human CGD
679 52 77.6 1966 7 ADJ70411 Adj70411 Human hea
680 52 77.6 2343 3 AAB12453 Aab12453 Human HNR
681 52 77.6 2440 8 ADQ89792 Adq89792 Antagonis

682 52 77.6 2440 9 ADW28417 Adw28417 Human nuc 755 51 76.1 242 4 AAM28950 Aam28950 Peptide #
683 52 77.6 2453 3 AAB12454 Aab12454 HNRCR pro 756 51 76.1 242 4 ABB30277 Abb30277 Peptide #
684 52 77.6 3119 2 AAW36888 Aaw36888 Mouse Hun 757 51 76.1 242 4 ABB20882 Abb20882 Protein #
685 52 77.6 3119 2 AAW44743 Aaw44743 Mouse hun 758 51 76.1 242 4 AAM68643 Aam68643 Human bon
686 52 77.6 3139 2 AAY08898 Aay08898 Human Hun 759 51 76.1 242 4 AAM56266 Aam56266 Human bra
687 52 77.6 3232 4 ABB69873 Abb69873 Drosophil 760 51 76.1 242 4 ABG50303 Abg50303 Human liv
688 51 76.1 10 2 AAR93316 Aar93316 ABL prote 761 51 76.1 242 4 AAM04184 Aam04184 Peptide #
689 51 76.1 29 1 AAP10194 Aap10194 Sequence 762 51 76.1 242 5 AAG38222 Abg38222 Human pep
690 51 76.1 29 1 AAP10194 Aap10194 Sequence 763 51 76.1 256 5 ABG33384 Abg33384 Human WT1
691 51 76.1 29 1 AAP71004 Aap71004 Sequence 764 51 76.1 256 7 ADB67537 Adb67537 Human WT1
692 51 76.1 29 1 AAP80038 Aap80038 Beta-huma 765 51 76.1 256 7 ADJ80770 Adj80770 Wilm's tu
693 51 76.1 29 1 AAP91844 Aap91844 Analogue 766 51 76.1 256 8 ADJ83690 Adj83690 Human WT1
694 51 76.1 29 2 AAW42211 Aaw42211 Peptide d 767 51 76.1 256 8 ADL57578 Adl57578 Human WT-
695 51 76.1 29 2 AAW69457 Aaw69457 HCG antig 768 51 76.1 256 8 ADO09133 Ado09133 Human WT1
696 51 76.1 29 2 AAW93442 Aaw93442 Human hCG 769 51 76.1 256 8 ADX68110 Adx68110 Plant ful
697 51 76.1 29 3 AAY87487 Aay87487 Human cho 770 51 76.1 266 7 ABM86694 Abm86694 Rice abio
698 51 76.1 29 3 AAB20551 Aab20551 Human cho 771 51 76.1 272 8 ADX90852 Adx90852 Plant ful
699 51 76.1 29 4 AAU01147 AAu01147 Structure 772 51 76.1 278 2 AAR79095 Aar79095 Rat Fas l
700 51 76.1 29 4 AAB48393 Aab48393 Human cho 773 51 76.1 278 2 AAW98069 Aaw98069 Rat Fas l
701 51 76.1 29 4 AAU02845 AAu02845 Human Cho 774 51 76.1 278 2 AAW95040 Aaw95040 Rat FasL
702 51 76.1 29 4 AAB04129 Aab04129 Peptide f 775 51 76.1 278 5 AAO19029 Aao19029 Human Fas
703 51 76.1 33 4 AAO04814 Aao04814 Human pol 776 51 76.1 278 7 ADD69944 Add69944 Rat Fas l
704 51 76.1 41 4 AAO02061 Aao02061 Human pol 777 51 76.1 278 8 ADP26982 Adp26982 Rat FasL
705 51 76.1 44 4 AAO04929 Aao04929 Human pol 778 51 76.1 278 8 ADU22867 Adu22867 Norway ra
706 51 76.1 69 7 ADC32722 Adc32722 Human nov 779 51 76.1 280 7 ADJ80896 Adj80896 Wilm's tu
707 51 76.1 70 3 AAG10803 Aag10803 Arabidops 780 51 76.1 280 8 ADJ83816 Adj83816 Human WT1
708 51 76.1 76 3 AAG10802 Aag10802 Arabidops 781 51 76.1 280 8 ADL57704 Adl57704 Human Wtl
709 51 76.1 77 4 AAM92774 Aam92774 Human dig 782 51 76.1 280 8 ADO09259 Ado09259 Human WT-
710 51 76.1 80 4 AAO04412 Aao04412 Human pol 783 51 76.1 281 7 ADJ80900 Adj80900 Wilm's tu
711 51 76.1 83 8 ADI45175 Adi45175 Rice isop 784 51 76.1 281 8 ADJ83820 Adj83820 Recombina
712 51 76.1 93 4 ABG04608 Abg04608 Novel hum 785 51 76.1 281 8 ADL57708 Adl57708 Human WT-
713 51 76.1 99 4 AAO02076 Aao02076 Human pol 786 51 76.1 281 8 ADO09263 Ado09263 Human WT-
714 51 76.1 99 5 ABG33386 Abg33386 Human Wtl 787 51 76.1 281 9 AEB55246 Aeb55246 Human b-z
715 51 76.1 99 7 ADB67544 Adb67544 Human pro 788 51 76.1 288 4 AAM80259 Aam80259 Human pro
716 51 76.1 99 7 ADJ80777 Adj80777 Wilm's tu 789 51 76.1 290 4 ABB68359 Abb68359 Drosophil
717 51 76.1 99 8 ADJ83697 Adj83697 Human Wtl 790 51 76.1 291 7 ADJ80899 Adj80899 Wilm's tu
718 51 76.1 99 8 ADL57585 Adl57585 Human Wtl 791 51 76.1 291 8 ADJ83819 Adj83819 Recombina
719 51 76.1 99 8 ADO09140 Ado09140 Human Wtl 792 51 76.1 291 8 ADL57707 Adl57707 Human WT-
720 51 76.1 109 4 AAO02138 Aao02138 Human pol 793 51 76.1 291 8 ADO09262 Ado09262 Human WT-
721 51 76.1 112 4 AAO02189 Aao02189 Human pol 794 51 76.1 292 7 ADJ80934 Adj80934 Wilm's tu
722 51 76.1 126 8 ADL06078 Adl06078 M. catarr 795 51 76.1 292 7 ADJ80905 Adj80905 Wilm's tu
723 51 76.1 128 4 AAO03145 Aao03145 Human pol 796 51 76.1 292 8 ADL57713 Adl57713 Human TAT
724 51 76.1 129 8 ADX94859 Adx94859 Plant ful 797 51 76.1 292 8 ADL57742 Adl57742 Human TAT
725 51 76.1 168 4 ABG13945 Abg13945 Novel hum 798 51 76.1 292 8 ADO09268 Ado09268 pstumpy-W
726 51 76.1 171 4 AAM17160 Aam17160 Peptide # 799 51 76.1 292 8 ADO09297 Ado09297 pstumpy-W
727 51 76.1 171 4 AAM15263 Aam15263 Peptide # 800 51 76.1 308 8 ADY04819 Ady04819 Plant ful
728 51 76.1 171 4 ABB36162 Abb36162 Peptide # 801 51 76.1 309 8 ADY04216 Ady04216 Plant ful
729 51 76.1 171 4 ABB34254 Abb34254 Peptide # 802 51 76.1 311 3 AAG17007 Aag17007 Arabidops
730 51 76.1 171 4 AAM29651 Aam29651 Peptide # 803 51 76.1 315 8 ADT56439 Adt56439 Plant pol
731 51 76.1 171 4 AAM27728 Aam27728 Peptide # 804 51 76.1 320 9 ADW17149 Adw17149 Eucalyptu
732 51 76.1 171 4 ABB30969 Abb30969 Peptide # 805 51 76.1 321 7 ADJ80914 Adj80914 Wilm's tu
733 51 76.1 171 4 ABB29093 Abb29093 Peptide # 806 51 76.1 321 7 ADJ80890 Adj80890 Wilm's tu
734 51 76.1 171 4 ABB19690 Abb19690 Protein # 807 51 76.1 321 8 ADJ83810 Adj83810 Human Wtl
735 51 76.1 171 4 ABB21546 Abb21546 Protein # 808 51 76.1 321 8 ADL57722 Adl57722 Human TAT
736 51 76.1 171 4 AAM67437 Aam67437 Human bon 809 51 76.1 321 8 ADO09253 Ado09253 Human Wtl
737 51 76.1 171 4 AAM69331 Aam69331 Human bon 810 51 76.1 321 8 ADO09277 Ado09277 Codon opt
738 51 76.1 171 4 AAM56944 Aam56944 Human bra 811 51 76.1 321 8 ADO09277 Ado09277 Codon opt
739 51 76.1 171 4 AAM55047 Aam55047 Human bra 812 51 76.1 327 7 ADC87451 Adc87451 Human GPC
740 51 76.1 171 4 ABG49093 Abg49093 Human liv 813 51 76.1 338 3 AAG17006 Aag17006 Arabidops
741 51 76.1 171 4 ABG51005 Abg51005 Human liv 814 51 76.1 347 3 AAB42695 Aab42695 Human ORF
742 51 76.1 171 4 AAM03008 Aam03008 Peptide # 815 51 76.1 347 7 ADG39809 Adg39809 Protein s
743 51 76.1 171 4 AAM04854 Aam04854 Peptide # 816 51 76.1 363 3 AAB07704 Aab07704 Protein e
744 51 76.1 171 5 ABB37060 Abg37060 Human pep 817 51 76.1 377 8 ADS98033 Ads98033 Protein f
745 51 76.1 171 5 ABG38947 Abg38947 Human pep 818 51 76.1 378 6 ABU07370 Abu07370 Human pro
746 51 76.1 175 4 AAO12806 Aao12806 Human pol 819 51 76.1 378 7 ADG39772 Adg39772 Human nov
747 51 76.1 186 8 ADK48382 Adk48382 Streptoco 820 51 76.1 391 3 AAY93904 Aay93904 Amino aci
748 51 76.1 199 7 ADC32888 Adc32888 Human nov 821 51 76.1 391 5 ABB92292 Abb92292 Herbicida
749 51 76.1 200 8 ADU01099 Adu01099 Human pro 822 51 76.1 391 8 ABB92292 Abb92292 Herbicida
750 51 76.1 200 8 ADU15481 Adu15481 Novel hum 823 51 76.1 404 8 ADX66693 Adx66693 Plant pol
751 51 76.1 210 2 AAR85066 Aar85066 Wilm's tu 824 51 76.1 406 4 ABG27250 Abg27250 Novel hum
752 51 76.1 210 2 AAW22883 Aaw22883 Wilm's tu 825 51 76.1 408 8 ADX67722 Adx67722 Plant ful
753 51 76.1 242 4 AAM16455 Aam16455 Peptide # 826 51 76.1 409 3 AAB07702 Aab07702 Protein e
754 51 76.1 242 4 ABB35443 Abb35443 Peptide # 827 51 76.1 410 5 ABG33382 Abg33382 Human Trx

828 51 76.1 410 7 ADB67535 Adb67535 Human WT1
829 51 76.1 410 7 ADJ80768 Adj80768 Wilm's tu
830 51 76.1 410 8 ADJ83688 Adj83688 Human WT1
831 51 76.1 410 8 ADL57576 Adl57576 Human Trx
832 51 76.1 410 8 ADO09131 Ado09131 Human Trx
833 51 76.1 417 5 ABG333399 Abg333399 Human WT1
834 51 76.1 417 7 ADB67609 Adb67609 Human WT1
835 51 76.1 417 7 ADJ80842 Adj80842 Wilm's tu
836 51 76.1 417 8 ADJ83762 Adj83762 Human WT1
837 51 76.1 417 8 ADL57650 Adl57650 Human WT1
838 51 76.1 417 8 ADO09205 Ado09205 Mutant hu
839 51 76.1 420 5 ABG333393 Abg333393 Human Ral
840 51 76.1 420 7 ADB67595 Adb67595 Human Ral
841 51 76.1 420 7 ADJ80828 Adj80828 Wilm's tu
842 51 76.1 420 8 ADJ83748 Adj83748 Human WT-
843 51 76.1 420 8 ADL57636 Adl57636 Human Ral
844 51 76.1 420 8 ADO09191 Ado09191 Human WT1
845 51 76.1 429 2 AAR85065 Aar85065 Wilm's tu
846 51 76.1 429 5 ABG333400 Abg333400 Human WT1
847 51 76.1 429 7 ADB67610 Adb67610 Human WT1
848 51 76.1 429 7 ADJ80843 Adj80843 Wilm's tu
849 51 76.1 429 8 ADJ83763 Adj83763 Human WT1
850 51 76.1 429 8 ADL57651 Adl57651 Human WT1
851 51 76.1 429 8 ADO09206 Ado09206 Human WT1
852 51 76.1 438 7 ADI21633 Adi21633 Novel hum
853 51 76.1 440 7 ABM90225 Abm90225 Rice abio
854 51 76.1 449 2 AAW47176 Aaw47176 Wilm's tu
855 51 76.1 449 3 AAY98805 Aay98805 Mouse WT1
856 51 76.1 449 3 AAY98804 Aay98804 Human WT1
857 51 76.1 449 3 AAY80196 Aay80196 Mouse WT1
858 51 76.1 449 3 AAY80197 Aay80197 Human WT1
859 51 76.1 449 4 AAG62136 Aag62136 Human WT1
860 51 76.1 449 4 AAG62137 Aag62137 Mouse WT1
861 51 76.1 449 4 AAU68904 Aau68904 Mouse Wil
862 51 76.1 449 4 AAU68903 Aau68903 Human Wil
863 51 76.1 449 5 ABG31644 Abg31644 Mouse Wil
864 51 76.1 449 5 ABG333374 Abg333374 Mouse WT1
865 51 76.1 449 5 ABG333373 Abg333373 Human WT1
866 51 76.1 449 5 ABG333396 Abg333396 Human WT1
867 51 76.1 449 5 AAG78446 Aag78446 Predicted
868 51 76.1 449 5 AAG78444 Aag78444 Murine WT
869 51 76.1 449 6 AB899880 Abb99880 Mouse Wil
870 51 76.1 449 6 ABG71423 Abg71423 Human Wil
871 51 76.1 449 6 ABG71421 Abg71421 Murine Wi
872 51 76.1 449 6 ABR44357 Abr44357 Human pro
873 51 76.1 449 6 ABR38970 Abr38970 Human tum
874 51 76.1 449 6 ABR38969 Abr38969 Mouse tum
875 51 76.1 449 6 ABR44448 Abr44448 Wilm's tu
876 51 76.1 449 7 ADB67606 Adb67606 Human Wil
877 51 76.1 449 7 ADB67522 Adb67522 Mouse Wil
878 51 76.1 449 7 ADB67521 Adb67521 Human Wil
879 51 76.1 449 7 ADJ80755 Adj80755 Wilm's tu
880 51 76.1 449 7 ADJ80839 Adj80839 Wilm's tu
881 51 76.1 449 7 ADJ80754 Adj80754 Wilm's tu
882 51 76.1 449 8 AAO24365 Aao24365 HLA-A24-r
883 51 76.1 449 8 ADJ83674 Adj83674 Human WT1
884 51 76.1 449 8 ADJ83759 Adj83759 Human WT1
885 51 76.1 449 8 ADJ83675 Adj83675 Murine WT
886 51 76.1 449 8 ADL57647 Adl57647 Human Wil
887 51 76.1 449 8 ADL57562 Adl57562 Human Wil
888 51 76.1 449 8 ADL57563 Adl57563 Mouse Wil
889 51 76.1 449 8 ADN00725 Adn00725 Human WT1
890 51 76.1 449 8 ADO09117 Ado09117 Human WT1
891 51 76.1 449 8 ADO09118 Ado09118 Mouse WT1
892 51 76.1 449 8 ADO09202 Ado09202 Human WT1
893 51 76.1 449 8 ADO43852 Ado43852 Amino aci
894 51 76.1 449 8 ADQ79585 Adq79585 Human Wil
895 51 76.1 449 9 ADW72806 Adw72806 Human WT1
896 51 76.1 449 9 AEA15654 Aea15654 Human WT1
897 51 76.1 453 7 ADB64865 Adb64865 Human pro
898 51 76.1 467 4 ABG27768 Abg27768 Novel hum
899 51 76.1 469 7 ADJ80889 Adj80889 Wilm's tu
900 51 76.1 469 8 ADJ83809 Adj83809 Human WT1

901 76.1 469 8 ADL57697
902 76.1 469 8 ADO09252
903 76.1 495 5 ABG33401
904 76.1 495 7 ADB67611
905 76.1 495 7 ADJ80844
906 76.1 495 8 ADJ83764
907 76.1 495 8 ADL57652
908 76.1 495 8 ADO09207
909 76.1 498 4 AAB73400
910 76.1 498 8 ABM80048
911 76.1 498 9 AEB56517
912 76.1 502 8 ADP12999
913 76.1 504 5 ABG33402
914 76.1 504 7 ADB67612
915 76.1 504 7 ADJ80845
916 76.1 504 8 ADJ83765
917 76.1 504 8 ADL57653
918 76.1 504 8 ADO09208
919 76.1 509 8 ADI42004
920 76.1 524 7 ABM89472
921 76.1 568 5 ABG33392
922 76.1 568 7 ADB67594
923 76.1 568 7 ADJ80827
924 76.1 568 8 ADJ83747
925 76.1 568 8 ADL57635
926 76.1 568 8 ADO09190
927 76.1 577 4 AAO22140
928 76.1 579 8 ADR66346
929 76.1 579 8 ADR66688
930 76.1 591 5 ABG33383
931 76.1 591 7 ADB67536
932 76.1 591 7 ADJ80769
933 76.1 591 8 ADJ83689
934 76.1 591 8 ADL57577
935 76.1 591 8 ADO09132
936 76.1 592 2 AAW82661
937 76.1 598 4 ABG14000
938 76.1 598 7 ADF60423
939 76.1 607 7 ADE08437
940 76.1 610 5 AAU76338
941 76.1 615 3 AAG46152
942 76.1 634 7 ADM04655
943 76.1 642 7 ADC94190
944 76.1 647 8 ADU01120
945 76.1 647 8 ADU15502
946 76.1 652 8 ADX96479
947 76.1 679 3 AAG46151
948 76.1 691 3 AAG46150
949 76.1 692 4 AAG62154
950 76.1 697 8 ADX97270
951 76.1 729 8 ADY07005
952 76.1 814 5 ABB92685
953 76.1 857 4 ABB69603
954 76.1 862 8 ADQ66190
955 76.1 883 7 ADM87622
956 76.1 963 7 ABM89266
957 76.1 1002 8 ADO01055
958 76.1 1006 8 ADU20559
959 76.1 1006 8 ADU20561
960 76.1 1086 4 ABB69356
961 76.1 1228 3 AAG30727
962 76.1 1230 3 AAG30726
963 76.1 1248 3 AAG30725
964 76.1 1413 5 AAE21729
965 76.1 1431 9 ADV97898
966 76.1 1460 7 ADC59465
967 76.1 1460 8 ADU04632
968 76.1 1518 6 ABJ18375
969 76.1 1737 8 ADO01057
970 76.1 1772 3 AAG52485
971 76.1 1774 5 ABG69800
972 76.1 1776 3 AAG52484
973 76.1 1787 3 AAG52483

Adl57697 Human Wil
Ado09252 Human WT-
Abg33401 Human WT1
Adb67611 Human WT1
Adj80844 Wilm's tu
Adj83764 Human WT-
Adl57652 Human LAM
Ado09207 Human LAM
Aab73400 Human WAV
Abm80048 Tumour-as
Aeb56517 Radiochem
Adp12999 Protein e
Abg33402 Human WT1
Adb67612 Human WT1
Adj80845 Wilm's tu
Adj83765 Human WT-
Adl57653 Human ubi
Ado09208 Human ubi
Adi42004 Plant tra
Abm89472 Rice abio
Abg33392 Human Ral
Adb67594 Human WT1
Adj80827 Wilm's tu
Adj83747 Human WT-
Adl57635 Human Ral
Ado09190 Human WT1
Aao22140 Ral2-WT1
Adr66346 Human pro
Adr66688 Human pro
Abg33383 Human Trx
Adb67536 Human WT1
Adj80769 Wilm's tu
Adj83689 Human WT1
Adl57577 Human Trx
Ado09132 Human Trx
Aaw82661 A. thalia
Abg14000 Novel hum
Adf60423 Human con
Ade08437 Novel pro
Aau76338 A. thalia
Aag46152 Arabidops
Adm04655 Human pro
Adc94190 E. faeciu
Adu01120 Human pro
Adu15502 Novel hum
Adx96479 Plant ful
Aag46151 Arabidops
Aag46150 Arabidops
Aag62154 Human WT1
Adx97270 Plant ful
Ady07005 Plant ful
Abb92685 Herbicida
Abb69603 Drosophil
Adq66190 Novel hum
Adm87622 Human EST
Abm89266 Rice abio
Ado01055 Mouse hom
Adu20559 A. thalia
Adu20561 A. thalia
Abb69356 Drosophil
Aag30727 Arabidops
Aag30726 Arabidops
Aag30725 Arabidops
Aae21729 Human PKI
Adv97898 Murine pr
Adc59465 Novel hum
Adu04632 Human KIA
Abj18375 Breast sp
Ado01057 Human hom
Aag52485 Arabidops
Abg69800 Human REM
Aag52484 Arabidops
Aag52483 Arabidops

974 51 76.1 1931 4 ABB66948 Abb66948 Drosophil
975 51 76.1 2778 7 ADC51660 Adc51660 Human MEG
976 51 76.1 2789 7 ADG75733 Adg75733 Human pro
977 51 76.1 2854 5 ADH48732 Nov7 prot
978 51 76.1 3051 6 ABR58308 BCU0947 p
979 51 76.1 3302 8 ADL26831 Rat L-NAM
980 51 76.1 3586 7 ADK40970 Nov1 hum
981 51 76.1 3586 8 ADR15689 Kinase 72
982 50.5 75.4 66 4 ABG16322 Nov1 hum
983 50.5 75.4 150 4 ABG08615 Nov1 hum
984 50 74.6 10 4 AAU09969 Proline 1
985 50 74.6 11 2 AAW98980 FAM- (Pro)
986 50 74.6 13 2 AAW98981 FAM- (Pro)
987 50 74.6 13 4 AAB30652 Trypsin m
988 50 74.6 15 2 AAW39030 Peptide r
989 50 74.6 15 4 AAB30653 Trypsin m
990 50 74.6 16 6 ABR62393 Synthetic
991 50 74.6 18 2 AAW38907 Peptide r
992 50 74.6 18 2 AAW38974 Peptide r
993 50 74.6 18 4 AAB30654 Trypsin m
994 50 74.6 19 2 AAW39009 Peptide r
995 50 74.6 19 4 AAB30655 Trypsin m
996 50 74.6 20 2 AAW38962 Peptide r
997 50 74.6 20 2 AAW38971 Peptide r
998 50 74.6 20 6 ABP59404 Self-asse
999 50 74.6 20 6 ABP59405 Self-asse
1000 50 74.6 21 4 AAB30656 Trypsin m

ALIGNMENTS

RESULT 1
AAR93314
ID AAR93314 standard; peptide; 10 AA.
XX
AC AAR93314;
XX
DT 24-APR-1996 (first entry)
DE ABL protein tyrosine kinase derived peptide #1.
XX
KW SH3 ligand; SH3 binding agent; biased phage library;
KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
KW p67; complex; chronic myelogenous leukaemia; cancer.
XX
OS Synthetic.
XX
PN WO9524419-A1.
XX
PD 14-SEP-1995.
XX
PF 13-MAR-1995; 95WO-US003208.
XX
PR 11-MAR-1994; 94US-00209835.
XX 06-JAN-1995; 95US-00369832.
PA (ARIA-) ARIAD PHARM INC.
XX
PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
XX
DR WPI; 1995-328231/42.
XX
PT Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.
XX
PS Disclosure; Fig 1; 74pp; English.
XX
CC The sequences given in AAR93272-342 represent peptides which are SH3
CC ligands/SH3 binding agents. They represent a biased phage library which
CC comprises six random amino acids flanking the tetrapeptide -PPIP which

CC was identified as a recognition sequence for the src SH3 domain. These
CC sequences were identified using the method of the invention. The method
CC comprises contacting the SH3 domain with a mixture of peptides under
CC conditions permitting a ligand to bind to an SH3 domain to form a
CC complex. Any unbound peptides are removed and the complexed peptide
CC ligands are dissociated from the complexes. The selected peptides are
CC enriched by re-contacting them with the SH3 domain and then candidates
CC which bind to the SH3 domain are detected. The isolated SH3 binding
CC peptides may be used in the diagnosis, prevention and treatment of
CC conditions or diseases resulting from cellular processes mediated by an
CC SH3-based interaction. Such diseases include Paget's disease. Other
CC conditions treatable with these peptides include restenosis, rheumatoid
CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
CC p47 and p67 complex is implicated, etc
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 67; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
Db 1 PPPYPPPPPIP 10
| | | | | | | | | |
RESULT 2
ADA08234
ID ADA08234 standard; peptide; 10 AA.
XX
AC ADA08234;
XX
DT 06-NOV-2003 (first entry)
DE Human Abl-SH3 related peptide SEQ ID NO:138.
XX
KW identification; interacting protein; plurality; protein display library;
KW binding protein; detection; specific binding;
KW target-assisted iterative screening; protein-protein interaction.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO2003029821-A1.
XX
PD 10-APR-2003.
XX
PF 01-OCT-2002; 2002WO-US031349.
XX
PR 01-OCT-2001; 2001US-0326566P.
XX
PA (BUCK-) BUCK INST.
XX
PI Kourakine A, Bredesen D;
XX
DR WPI; 2003-403031/38.
XX
PT Identifying interacting proteins from a protein display library having the
PT potential binding proteins comprises contacting target proteins with the
PT library and detecting binding of potential binding proteins with the
PT target proteins.
XX
PS Example 1; Page 45; 61pp; English.
XX
CC The present invention describes a method for identifying interacting
CC proteins from a plurality of potentially-interacting proteins comprising
CC contacting one or more target proteins with a protein display library
CC having a plurality of potential binding proteins and detecting specific
CC binding of potential binding proteins with the target proteins. The
CC method comprises: (a) contacting one or more target proteins with a
CC protein display library comprising a plurality of potential binding
CC proteins for one or more target proteins; (b) selecting members of the
CC protein display library that bind to the target proteins to provide a

CC preselected set of potential binding proteins; (c) separating the members
CC of the preselected set of potential binding proteins from the bound
CC target protein and immobilising the members on a solid support such that
CC the members are spatially addressable; (d) contacting the members of the
CC preselected set of potential binding proteins with the target proteins;
CC and (e) detecting specific binding of members of the preselected set of
CC potential binding proteins with the target proteins, where binding of a
CC member of the preselected set with a target protein indicates that the
CC member and the target protein are interacting proteins. Also described is
CC a kit for carrying out the method described above, comprising a protein
CC display library and instructional materials providing protocols for the
CC above method. The target-assisted iterative screening method is useful
CC for determining large numbers of interactions (e.g. protein-protein
CC interactions) between members of a library and various targets. The
CC present sequence represents a peptide, which is used in the
CC exemplification of the present invention.

XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 67; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
Db 1 PPPYPPPPPIP 10

RESULT 3
AAR93369
ID AAR93369 standard; peptide; 12 AA.
XX
AC AAR93369;
XX
DT 24-APR-1996 (first entry)
XX
DE ABL protein tyrosine kinase derived peptide #1.
XX
KW SH3 ligand; SH3 binding agent; biased phage library;
KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
KW p67; complex; chronic myelogenous leukaemia; cancer.

XX Synthetic.
XX WO9524419-A1.
XX
PD 14-SEP-1995.
XX
PF 13-MAR-1995; 95WO-US003208.
XX
PR 11-MAR-1994; 94US-00209835.
PR 06-JAN-1995; 95US-00369832.
XX
PA (ARIA-) ARIAD PHARM INC.
XX
PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
XX WPI; 1995-328231/42.
DR

XX Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.

XX Disclosure; Fig 2; 74pp; English.
XX
CC The sequences given in AAR93369-76 represent peptides which are SH3
CC ligands/SH3 binding agents. They represent a biased phage library which
CC comprises six random amino acids flanking the hexapeptide PPPYPP- which
CC was identified as a recognition sequence for the src SH3 domain. These
CC sequences were identified using the method of the invention. The method
CC comprises contacting the SH3 domain with a mixture of peptides under
CC conditions permitting a ligand to bind to an SH3 domain to form a

CC complex. Any unbound peptides are removed and the complexed peptide
CC ligands are dissociated from the complexes. The selected peptides are
CC enriched by re-contacting them with the SH3 domain and then candidates
CC which bind to the SH3 domain are detected. The isolated SH3 binding
CC peptides may be used in the diagnosis, prevention and treatment of
CC conditions or diseases resulting from cellular processes mediated by an
CC SH3-based interaction. Such diseases include Paget's disease. Other
CC conditions treatable with these peptides include restenosis, rheumatoid
CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
CC p47 and p67 complex is implicated, etc

XX Sequence 12 AA;

Query Match 100.0%; Score 67; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.56;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
Db 1 PPPYPPPPPIP 10

RESULT 4
AAR93371
ID AAR93371 standard; peptide; 12 AA.
XX
AC AAR93371;
XX
DT 24-APR-1996 (first entry)
XX
DE ABL protein tyrosine kinase derived peptide #3.
XX
KW SH3 ligand; SH3 binding agent; biased phage library;
KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
KW p67; complex; chronic myelogenous leukaemia; cancer.

XX Synthetic.
XX WO9524419-A1.
XX
PD 14-SEP-1995.
XX
PF 13-MAR-1995; 95WO-US003208.
XX
PR 11-MAR-1994; 94US-00209835.
PR 06-JAN-1995; 95US-00369832.
XX
PA (ARIA-) ARIAD PHARM INC.
XX
PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
XX WPI; 1995-328231/42.

XX Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.

XX Disclosure; Fig 2; 74pp; English.

XX
CC The sequences given in AAR93369-76 represent peptides which are SH3
CC ligands/SH3 binding agents. They represent a biased phage library which
CC comprises six random amino acids flanking the hexapeptide PPPYPP- which
CC was identified as a recognition sequence for the src SH3 domain. These
CC sequences were identified using the method of the invention. The method
CC comprises contacting the SH3 domain with a mixture of peptides under
CC conditions permitting a ligand to bind to an SH3 domain to form a
CC complex. Any unbound peptides are removed and the complexed peptide
CC ligands are dissociated from the complexes. The selected peptides are
CC enriched by re-contacting them with the SH3 domain and then candidates
CC which bind to the SH3 domain are detected. The isolated SH3 binding
CC peptides may be used in the diagnosis, prevention and treatment of
CC conditions or diseases resulting from cellular processes mediated by an

CC SH3-based interaction. Such diseases include Paget's disease. Other
CC conditons treatable with these peptides include restenosis, rheumatoid
CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
CC p47 and p67 complex is implicated, etc

XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 67; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.56;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
| | | | | | | | | |
Db 1 PPPYPPPPPIP 10

RESULT 5
AAB17256
ID AAB17256 standard; peptide; 12 AA.
XX
AC AAB17256;
XX
DT 31-OCT-2000 (first entry)
XX
DE SH3 antagonist peptide sequence SEQ ID NO:312.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 39; Page 304; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the

CC exemplification of the present invention
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 67; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.56;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
| | | | | | | | | |
Db 1 PPPYPPPPPIP 10

RESULT 6
ABB73249
ID ABB73249 standard; peptide; 12 AA.
XX
AC ABB73249;
XX
DT 05-APR-2002 (first entry)
XX
DE Src homology3 (SH3) antagonist peptide SEQ ID NO:312.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 39; Page 55; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low

CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 67; DB 5; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.56;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 1 PPPYPPPPPIP 10

RESULT 7
ADJ73403
ID ADJ73403 standard; peptide; 12 AA.
XX
AC ADJ73403;
XX
DT 06-MAY-2004 (first entry)
XX
DE SH3 antagonist peptide sequence SeqID 858.
XX
KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW SH3.
XX
OS Synthetic.
XX
PN WO2003084477-A2.
XX
PD 16-OCT-2003.
XX
PF 24-MAR-2003; 2003WO-US009139.
XX
PR 29-MAR-2002; 2002US-0368791P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Scallion BJ, Ghrayeb J;
XX
DR WPI; 2003-804237/75.
XX
PT New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.
XX
PS Disclosure; SEQ ID NO 858; 97pp; English.
XX
CC This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/ or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is an SH3 antagonist peptide sequence used to make a

CC mimetibody of the invention.
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 67; DB 7; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.56;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 1 PPPYPPPPPIP 10

RESULT 8
ADJ53037
ID ADJ53037 standard; peptide; 12 AA.
XX
AC ADJ53037;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID858.
XX
KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX
OS Unidentified.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1..12 /notes= "All Xaa's in this sequence are unidentified amino
FT acids"
XX
PN WO2004002417-A2.
XX
PD 08-JAN-2004.
XX
PF 27-JUN-2003; 2003WO-US020347.
XX
PR 28-JUN-2002; 2002US-0392431P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082870/08.
XX
PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX
PS Claim 3; SEQ ID NO 858; 129pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous

CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 67; DB 8; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.56;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPYPPPPIP 10
Db 1 PPPYPPPPIP 10

RESULT 9
ADJ51998
ID ADJ51998 standard; peptide; 12 AA.
XX
AC ADJ51998;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID858.
XX
KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmologic; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.
XX
OS Unidentified.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1. .12 /notes= "All Xaa's in this sequence are unidentified amino
FT acids"
XX
PN WO2004002424-A2.
XX
PD 08-JAN-2004.
XX
PF 30-JUN-2003; 2003WO-US020495.
XX
PR 28-JUN-2002; 2002US-0392431P.
PR 19-SEP-2002; 2002US-0412144P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutolowski KA;
XX
DR WPI; 2004-082872/08.
XX
PT New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.
XX
PS Claim 15; SEQ ID NO 858; 123pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be

CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC antiinflammatory, neuroleptic, ophthalmologic, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CH1 deleted mimetibody, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetibody of the invention.
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 67; DB 8; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.56;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPYPPPPIP 10
Db 1 PPPYPPPPIP 10

RESULT 10
AAR93373
ID AAR93373 standard; peptide; 12 AA.
XX
AC AAR93373;
XX
DT 24-APR-1996 (first entry)
XX
DE ABL protein tyrosine kinase derived peptide #5.
XX
KW SH3 ligand; SH3 binding agent; biased phage library;
KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
KW p67; complex; chronic myelogenous leukaemia; cancer.
XX
OS Synthetic.
XX
PN WO9524419-A1.
XX
PD 14-SEP-1995.
XX
PF 13-MAR-1995; 95WO-US003208.
XX
PR 11-MAR-1994; 94US-00209835.
PR 06-JAN-1995; 95US-00369832.
XX
PA (ARIA-) ARIAD PHARM INC.
XX
PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
XX
DR WPI; 1995-328231/42.
XX
PT Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.
XX
PS Disclosure; Fig 2; 74pp; English.
XX
CC The sequences given in AAR93369-76 represent peptides which are SH3
CC ligands/SH3 binding agents. They represent a biased phage library which
CC comprises six random amino acids flanking the hexapeptide PPPYPP- which
CC was identified as a recognition sequence for the src SH3 domain. These
CC sequences were identified using the method of the invention. The method
CC comprises contacting the SH3 domain with a mixture of peptides under

CC conditions permitting a ligand to bind to an SH3 domain to form a
CC complex. Any unbound peptides are removed and the complexed peptide
CC ligands are dissociated from the complexes. The selected peptides are
CC enriched by re-contacting them with the SH3 domain and then candidates
CC which bind to the SH3 domain are detected. The isolated SH3 binding
CC peptides may be used in the diagnosis, prevention and treatment of
CC conditions or diseases resulting from cellular processes mediated by an
CC SH3-based interaction. Such diseases include Paget's disease. Other
CC conditions treatable with these peptides include restenosis, rheumatoid
CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
CC p47 and p67 complex is implicated, etc
XX
SQ Sequence 12 AA;

Query Match 98.5%; Score 66; DB 2; Length 12;
Best Local Similarity 90.0%; Pred. No. 0.71;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
Db 1 PPPYPPPPVP 10

RESULT 11
AAR933372
ID AAR933372 standard; peptide; 12 AA.
XX
AC AAR933372;
XX
DT 24-APR-1996 (first entry)
XX
DE ABL protein tyrosine kinase derived peptide #4.
XX
KW SH3 ligand; SH3 binding agent; biased phage library;
KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
KW p67; complex; chronic myelogenous leukaemia; cancer.
XX
OS Synthetic.
XX
PN WO9524419-A1.
XX
PD 14-SEP-1995.
XX
PF 13-MAR-1995; 95WO-US003208.
XX
PR 11-MAR-1994; 94US-00209835.
PR 06-JAN-1995; 95US-00369832.
XX
PA (ARIA-) ARIAD PHARM INC.
XX
PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
XX
DR WPI; 1995-328231/42.
XX
PT Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.
XX
PS Disclosure; Fig 2; 74pp; English.
XX
CC The sequences given in AAR93369-76 represent peptides which are SH3
CC ligands/SH3 binding agents. They represent a biased phage library which
CC comprises six random amino acids flanking the hexapeptide PPPYPP- which
CC was identified as a recognition sequence for the src SH3 domain. These
CC sequences were identified using the method of the invention. The method
CC comprises contacting the SH3 domain with a mixture of peptides under
CC conditions permitting a ligand to bind to an SH3 domain to form a
CC complex. Any unbound peptides are removed and the complexed peptide
CC ligands are dissociated from the complexes. The selected peptides are
CC enriched by re-contacting them with the SH3 domain and then candidates
CC which bind to the SH3 domain are detected. The isolated SH3 binding
CC peptides may be used in the diagnosis, prevention and treatment of

CC conditions or diseases resulting from cellular processes mediated by an
CC SH3-based interaction. Such diseases include Paget's disease. Other
CC conditions treatable with these peptides include restenosis, rheumatoid
CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
CC p47 and p67 complex is implicated, etc
XX
SQ Sequence 12 AA;

Query Match 98.5%; Score 66; DB 2; Length 12;
Best Local Similarity 90.0%; Pred. No. 0.71;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
Db 1 PPPYPPPPVP 10

RESULT 12
AAR933370
ID AAR933370 standard; peptide; 12 AA.
XX
AC AAR933370;
XX
DT 24-APR-1996 (first entry)
XX
DE ABL protein tyrosine kinase derived peptide #2.
XX
KW SH3 ligand; SH3 binding agent; biased phage library;
KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
KW p67; complex; chronic myelogenous leukaemia; cancer.
XX
OS Synthetic.
XX
PN WO9524419-A1.
XX
PD 14-SEP-1995.
XX
PF 13-MAR-1995; 95WO-US003208.
XX
PR 11-MAR-1994; 94US-00209835.
PR 06-JAN-1995; 95US-00369832.
XX
PA (ARIA-) ARIAD PHARM INC.
XX
PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
XX
DR WPI; 1995-328231/42.
XX
PT Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.
XX
PS Disclosure; Fig 2; 74pp; English.
XX
CC The sequences given in AAR93369-76 represent peptides which are SH3
CC ligands/SH3 binding agents. They represent a biased phage library which
CC comprises six random amino acids flanking the hexapeptide PPPYPP- which
CC was identified as a recognition sequence for the src SH3 domain. These
CC sequences were identified using the method of the invention. The method
CC comprises contacting the SH3 domain with a mixture of peptides under
CC conditions permitting a ligand to bind to an SH3 domain to form a
CC complex. Any unbound peptides are removed and the complexed peptide
CC ligands are dissociated from the complexes. The selected peptides are
CC enriched by re-contacting them with the SH3 domain and then candidates
CC which bind to the SH3 domain are detected. The isolated SH3 binding
CC peptides may be used in the diagnosis, prevention and treatment of
CC conditions or diseases resulting from cellular processes mediated by an
CC SH3-based interaction. Such diseases include Paget's disease. Other
CC conditions treatable with these peptides include restenosis, rheumatoid
CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
CC p47 and p67 complex is implicated, etc
XX

SQ Sequence 12 AA;
Query Match 98.5%; Score 66; DB 2; Length 12;
Best Local Similarity 90.0%; Pred. No. 0.71;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 1 PPPYPPPPVP 10

RESULT 13
AAB17257
ID AAB17257 standard; peptide; 12 AA.
XX
AC AAB17257;
XX
DT 31-OCT-2000 (first entry)
XX
DE SH3 antagonist peptide sequence SEQ ID NO:313.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 39; Page 304; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 12 AA;
Query Match 98.5%; Score 66; DB 3; Length 12;

Best Local Similarity 90.0%; Pred. No. 0.71;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 1 PPPYPPPPVP 10

RESULT 14
ABB73250
ID ABB73250 standard; peptide; 12 AA.
XX
AC ABB73250;
XX
DT 05-APR-2002 (first entry)
XX
DE Src homology3 (SH3) antagonist peptide SEQ ID NO:313.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 39; Page 55; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,

CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention

XX SQ Sequence 12 AA;

Query Match 98.5%; Score 66; DB 5; Length 12;
Best Local Similarity 90.0%; Pred. No. 0.71;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
Db 1 PPPYPPPPVP 10

RESULT 15
ADJ73404
ID ADJ73404 standard; peptide; 12 AA.
XX
AC ADJ73404;
XX
DT 06-MAY-2004 (first entry)
XX
DE SH3 antagonist peptide sequence SeqID 859.
XX
KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW SH3.
XX
OS Synthetic.
XX
PN WO2003084477-A2.
XX
PD 16-OCT-2003.
XX
PF 24-MAR-2003; 2003WO-US009139.
XX
PR 29-MAR-2002; 2002US-0368791P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Scallion BJ, Ghrayeb J;
XX
DR WPI; 2003-804237/75.
XX
PT New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.
XX
PS Disclosure; SEQ ID NO 859; 97pp; English.
XX
CC This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/ or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is an SH3 antagonist peptide sequence used to make a
CC mimetibody of the invention.

XX SQ Sequence 12 AA;

Query Match 98.5%; Score 66; DB 7; Length 12;

Best Local Similarity 90.0%; Pred. No. 0.71;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
Db 1 PPPYPPPPVP 10

RESULT 16
ADJ53038
ID ADJ53038 standard; peptide; 12 AA.
XX
AC ADJ53038;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID859.
XX
KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX
OS Unidentified.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1. .12
FT /note= "All Xaa's in this sequence are unidentified amino
FT acids"
XX
PN WO2004002417-A2.
XX
PD 08-JAN-2004.
XX
PF 27-JUN-2003; 2003WO-US020347.
XX
PR 28-JUN-2002; 2002US-0392431P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082870/08.
XX
PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX
PS Claim 3; SEQ ID NO 859; 129pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.

XX SQ Sequence 12 AA;

Query Match 98.5%; Score 66; DB 8; Length 12;
Best Local Similarity 90.0%; Pred. No. 0.71;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
Db 1 PPPYPPPPVP 10

RESULT 17
ADJ51999
ID ADJ51999 standard; peptide; 12 AA.
XX
AC ADJ51999;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID859.
XX
KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunologic disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.
XX
OS Unidentified.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1..12
FT /note= "All Xaa's in this sequence are unidentified amino
FT acids"
XX
PN WO2004002424-A2.
XX
PD 08-JAN-2004.
XX
PF 30-JUN-2003; 2003WO-US020495.
XX
PR 28-JUN-2002; 2002US-0392431P.
PR 19-SEP-2002; 2002US-0412144P.
XX
PA (GENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082872/08.
XX
PT New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.
XX
PS Claim 15; SEQ ID NO 859; 123pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or

CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CH1 deleted mimetibody, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetibody of the invention.

XX SQ Sequence 12 AA;

Query Match 98.5%; Score 66; DB 8; Length 12;
Best Local Similarity 90.0%; Pred. No. 0.71;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
Db 1 PPPYPPPPVP 10

RESULT 18

AAR93315
ID AAR93315 standard; peptide; 10 AA.
XX
AC AAR93315;
XX
DT 24-APR-1996 (first entry)
XX
DE ABL protein tyrosine kinase derived peptide #2.
XX
KW SH3 ligand; SH3 binding agent; biased phage library;
KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
KW p67; complex; chronic myelogenous leukaemia; cancer.
XX
OS Synthetic.
XX
PN WO9524419-A1.
XX
PD 14-SEP-1995.
XX
PF 13-MAR-1995; 95WO-US003208.
XX
PR 11-MAR-1994; 94US-00209835.
PR 06-JAN-1995; 95US-00369832.
XX
PA (ARIA-) ARIAD PHARM INC.
XX
PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
XX
DR WPI; 1995-328231/42.
XX
PT Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.
XX
PS Disclosure; Fig 1; 74pp; English.
XX

CC The sequences given in AAR93272-342 represent peptides which are SH3
CC ligands/SH3 binding agents. They represent a biased phage library which
CC comprises six random amino acids flanking the tetrapeptide -PPIP which
CC was identified as a recognition sequence for the src SH3 domain. These
CC sequences were identified using the method of the invention. The method
CC comprises contacting the SH3 domain with a mixture of peptides under
CC conditions permitting a ligand to bind to an SH3 domain to form a
CC complex. Any unbound peptides are removed and the complexed peptide
CC ligands are dissociated from the complexes. The selected peptides are
CC enriched by re-contacting them with the SH3 domain and then candidates
CC which bind to the SH3 domain are detected. The isolated SH3 binding

CC peptides may be used in the diagnosis, prevention and treatment of
CC conditions or diseases resulting from cellular processes mediated by an
CC SH3-based interaction. Such diseases include Paget's disease. Other
CC conditions treatable with these peptides include restenosis, rheumatoid
CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
CC p47 and p67 complex is implicated, etc
XX
SQ Sequence 10 AA;

Query Match 89.6%; Score 60; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PYPYPPPIP 10
Db 2 PYPYPPPIP 10

RESULT 19
AAR93318
ID AAR93318 standard; peptide; 10 AA.
XX
AC AAR93318;
XX
DT 24-APR-1996 (first entry)
XX
DE ABL protein tyrosine kinase derived peptide #5.
XX
KW SH3 ligand; SH3 binding agent; biased phage library;
KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
KW p67; complex; chronic myelogenous leukaemia; cancer.
XX
OS Synthetic.
XX
PN WO9524419-A1.
XX
PD 14-SEP-1995.
XX
PF 13-MAR-1995; 95WO-US003208.
XX
PR 11-MAR-1994; 94US-00209835.
PR 06-JAN-1995; 95US-00369832.
XX
PA (ARIA-) ARIAD PHARM INC.
XX
PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
XX
DR WPI; 1995-328231/42.
XX
PT Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.
XX
PS Disclosure; Fig 1; 74pp; English.
XX
CC The sequences given in AAR93272-342 represent peptides which are SH3
CC ligands/SH3 binding agents. They represent a biased phage library which
CC comprises six random amino acids flanking the tetrapeptide -PPIP which
CC was identified as a recognition sequence for the src SH3 domain. These
CC sequences were identified using the method of the invention. The method
CC comprises contacting the SH3 domain with a mixture of peptides under
CC conditions permitting a ligand to bind to an SH3 domain to form a
CC complex. Any unbound peptides are removed and the complexed peptide
CC ligands are dissociated from the complexes. The selected peptides are
CC enriched by re-contacting them with the SH3 domain and then candidates
CC which bind to the SH3 domain are detected. The isolated SH3 binding
CC peptides may be used in the diagnosis, prevention and treatment of
CC conditions or diseases resulting from cellular processes mediated by an
CC SH3-based interaction. Such diseases include Paget's disease. Other
CC conditions treatable with these peptides include restenosis, rheumatoid
CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
CC p47 and p67 complex is implicated, etc

XX
SQ Sequence 10 AA;

Query Match 88.1%; Score 59; DB 2; Length 10;
Best Local Similarity 90.0%; Pred. No. 3.4;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PYPYPPPIP 10
Db 1 PPSYPPPIP 10

RESULT 20
AAR93319
ID AAR93319 standard; peptide; 10 AA.
XX
AC AAR93319;
XX
DT 24-APR-1996 (first entry)
XX
DE ABL protein tyrosine kinase derived peptide #6.
XX
KW SH3 ligand; SH3 binding agent; biased phage library;
KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
KW p67; complex; chronic myelogenous leukaemia; cancer.
XX
OS Synthetic.
XX
PN WO9524419-A1.
XX
PD 14-SEP-1995.
XX
PF 13-MAR-1995; 95WO-US003208.
XX
PR 11-MAR-1994; 94US-00209835.
PR 06-JAN-1995; 95US-00369832.
XX
PA (ARIA-) ARIAD PHARM INC.
XX
PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
XX
DR WPI; 1995-328231/42.
XX
PT Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.
XX
PS Disclosure; Fig 1; 74pp; English.
XX
CC The sequences given in AAR93272-342 represent peptides which are SH3
CC ligands/SH3 binding agents. They represent a biased phage library which
CC comprises six random amino acids flanking the tetrapeptide -PPIP which
CC was identified as a recognition sequence for the src SH3 domain. These
CC sequences were identified using the method of the invention. The method
CC comprises contacting the SH3 domain with a mixture of peptides under
CC conditions permitting a ligand to bind to an SH3 domain to form a
CC complex. Any unbound peptides are removed and the complexed peptide
CC ligands are dissociated from the complexes. The selected peptides are
CC enriched by re-contacting them with the SH3 domain and then candidates
CC which bind to the SH3 domain are detected. The isolated SH3 binding
CC peptides may be used in the diagnosis, prevention and treatment of
CC conditions or diseases resulting from cellular processes mediated by an
CC SH3-based interaction. Such diseases include Paget's disease. Other
CC conditions treatable with these peptides include restenosis, rheumatoid
CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
CC p47 and p67 complex is implicated, etc
XX
SQ Sequence 10 AA;

Query Match 88.1%; Score 59; DB 2; Length 10;
Best Local Similarity 90.0%; Pred. No. 3.4;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPYP PPPPIP 10
Db 1 PPAY PPPPIP 10

RESULT 21
AAR93374
ID AAR93374 standard; peptide; 11 AA.
XX
AC AAR93374;
XX
DT 24-APR-1996 (first entry)
XX
DE ABL protein tyrosine kinase derived peptide #6.
XX
KW SH3 ligand; SH3 binding agent; biased phage library;
KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
KW p67; complex; chronic myelogenous leukaemia; cancer.
XX
OS Synthetic.
XX
PN WO9524419-A1.
XX
PD 14-SEP-1995.
XX
PF 13-MAR-1995; 95WO-US003208.
XX
PR 11-MAR-1994; 94US-00209835.
PR 06-JAN-1995; 95US-00369832.
XX
PA (ARIA-) ARIAD PHARM INC.
XX
PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
XX WPI; 1995-328231/42.
XX
PT Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.
XX
PS Disclosure; Fig 2; 74pp; English.
XX
CC The sequences given in AAR93369-76 represent peptides which are SH3
CC ligands/SH3 binding agents. They represent a biased phage library which
CC comprises six random amino acids flanking the hexapeptide PPYP- which
CC was identified as a recognition sequence for the src SH3 domain. These
CC sequences were identified using the method of the invention. The method
CC comprises contacting the SH3 domain with a mixture of peptides under
CC conditions permitting a ligand to bind to an SH3 domain to form a
CC complex. Any unbound peptides are removed and the complexed peptide
CC ligands are dissociated from the complexes. The selected peptides are
CC enriched by re-contacting them with the SH3 domain and then candidates
CC which bind to the SH3 domain are detected. The isolated SH3 binding
CC peptides may be used in the diagnosis, prevention and treatment of
CC conditions or diseases resulting from cellular processes mediated by an
CC SH3-based interaction. Such diseases include Paget's disease. Other
CC conditions treatable with these peptides include restenosis, rheumatoid
CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
CC p47 and p67 complex is implicated, etc
XX
SQ Sequence 11 AA;
Query Match 88.1%; Score 59; DB 2; Length 11;
Best Local Similarity 88.9%; Pred. No. 3.6;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PPYP PPPPIP 10
Db 1 PPYP PPPPVP 9

RESULT 22
AAR93376
ID AAR93376 standard; peptide; 12 AA.
XX
AC AAR93376;
XX
DT 24-APR-1996 (first entry)
XX
DE ABL protein tyrosine kinase derived peptide #8.
XX
KW SH3 ligand; SH3 binding agent; biased phage library;
KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
KW p67; complex; chronic myelogenous leukaemia; cancer.
XX
OS Synthetic.
XX
PN WO9524419-A1.
XX
PD 14-SEP-1995.
XX
PF 13-MAR-1995; 95WO-US003208.
XX
PR 11-MAR-1994; 94US-00209835.
PR 06-JAN-1995; 95US-00369832.
XX
PA (ARIA-) ARIAD PHARM INC.
XX
PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
XX WPI; 1995-328231/42.
XX
PT Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.
XX
PS Disclosure; Fig 2; 74pp; English.
XX
CC The sequences given in AAR93369-76 represent peptides which are SH3
CC ligands/SH3 binding agents. They represent a biased phage library which
CC comprises six random amino acids flanking the hexapeptide PPYP- which
CC was identified as a recognition sequence for the src SH3 domain. These
CC sequences were identified using the method of the invention. The method
CC comprises contacting the SH3 domain with a mixture of peptides under
CC conditions permitting a ligand to bind to an SH3 domain to form a
CC complex. Any unbound peptides are removed and the complexed peptide
CC ligands are dissociated from the complexes. The selected peptides are
CC enriched by re-contacting them with the SH3 domain and then candidates
CC which bind to the SH3 domain are detected. The isolated SH3 binding
CC peptides may be used in the diagnosis, prevention and treatment of
CC conditions or diseases resulting from cellular processes mediated by an
CC SH3-based interaction. Such diseases include Paget's disease. Other
CC conditions treatable with these peptides include restenosis, rheumatoid
CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
CC p47 and p67 complex is implicated, etc
XX
SQ Sequence 12 AA;
Query Match 88.1%; Score 59; DB 2; Length 12;
Best Local Similarity 90.0%; Pred. No. 3.9;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPYP PPPPIP 10
Db 1 PPYP PPPDIP 10

RESULT 23
AAR93317
ID AAR93317 standard; peptide; 10 AA.
XX
AC AAR93317;
XX

DT 24-APR-1996 (first entry)
XX ABL protein tyrosine kinase derived peptide #4.
DE
XX SH3 ligand; SH3 binding agent; biased phage library;
KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
KW p67; complex; chronic myelogenous leukaemia; cancer.
XX
OS Synthetic.
XX
PN WO9524419-A1.
XX
PD 14-SEP-1995.
XX
PF 13-MAR-1995; 95WO-US003208.
XX
PR 11-MAR-1994; 94US-00209835.
PR 06-JAN-1995; 95US-00369832.
XX
PA (ARIA-) ARIAD PHARM INC.
XX
PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
XX WPI; 1995-328231/42.
DR
XX Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.
XX
PS Disclosure; Fig 1; 74pp; English.
XX
CC The sequences given in AAR93272-342 represent peptides which are SH3
CC ligands/SH3 binding agents. They represent a biased phage library which
CC comprises six random amino acids flanking the tetrapeptide -PPIP which
CC was identified as a recognition sequence for the src SH3 domain. These
CC sequences were identified using the method of the invention. The method
CC comprises contacting the SH3 domain with a mixture of peptides under
CC conditions permitting a ligand to bind to an SH3 domain to form a
CC complex. Any unbound peptides are removed and the complexed peptide
CC ligands are dissociated from the complexes. The selected peptides are
CC enriched by re-contacting them with the SH3 domain and then candidates
CC which bind to the SH3 domain are detected. The isolated SH3 binding
CC peptides may be used in the diagnosis, prevention and treatment of
CC conditions or diseases resulting from cellular processes mediated by an
CC SH3-based interaction. Such diseases include Paget's disease. Other
CC conditions treatable with these peptides include restenosis, rheumatoid
CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
CC p47 and p67 complex is implicated, etc
XX
SQ Sequence 10 AA;
Query Match 86.6%; Score 58; DB 2; Length 10;
Best Local Similarity 90.0%; Pred. No. 4.3;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 PPPYPPPPPIP 10
Db 1 PPYPYHPPPIP 10
RESULT 24
AAR93323
ID AAR93323 standard; peptide; 10 AA.
XX
AC AAR93323;
XX
DT 24-APR-1996 (first entry)
XX
DE ABL protein tyrosine kinase derived peptide #10.
XX
KW SH3 ligand; SH3 binding agent; biased phage library;
KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
KW

KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
KW p67; complex; chronic myelogenous leukaemia; cancer.
XX
OS Synthetic.
XX
PN WO9524419-A1.
XX
PD 14-SEP-1995.
XX
PF 13-MAR-1995; 95WO-US003208.
XX
PR 11-MAR-1994; 94US-00209835.
PR 06-JAN-1995; 95US-00369832.
XX
PA (ARIA-) ARIAD PHARM INC.
XX
PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
XX WPI; 1995-328231/42.
DR
XX Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.
XX
PS Disclosure; Fig 1; 74pp; English.
XX
CC The sequences given in AAR93272-342 represent peptides which are SH3
CC ligands/SH3 binding agents. They represent a biased phage library which
CC comprises six random amino acids flanking the tetrapeptide -PPIP which
CC was identified as a recognition sequence for the src SH3 domain. These
CC sequences were identified using the method of the invention. The method
CC comprises contacting the SH3 domain with a mixture of peptides under
CC conditions permitting a ligand to bind to an SH3 domain to form a
CC complex. Any unbound peptides are removed and the complexed peptide
CC ligands are dissociated from the complexes. The selected peptides are
CC enriched by re-contacting them with the SH3 domain and then candidates
CC which bind to the SH3 domain are detected. The isolated SH3 binding
CC peptides may be used in the diagnosis, prevention and treatment of
CC conditions or diseases resulting from cellular processes mediated by an
CC SH3-based interaction. Such diseases include Paget's disease. Other
CC conditions treatable with these peptides include restenosis, rheumatoid
CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
CC p47 and p67 complex is implicated, etc
XX
SQ Sequence 10 AA;
Query Match 86.6%; Score 58; DB 2; Length 10;
Best Local Similarity 90.0%; Pred. No. 4.3;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 PPPYPPPPPIP 10
Db 1 PPYPYHPPPIP 10
RESULT 25
AAB98984
ID AAB98984 standard; peptide; 10 AA.
XX
AC AAB98984;
XX
DT 20-AUG-2001 (first entry)
XX
DE Yeast ORF YDL117W SH3 domain peptide ligand.
XX
KW Yeast; ORF YDL117W; protein 3-dimensional modelling; 3-d modelling;
KW protein structure analysis; ligand.
XX
OS Unidentified.
XX
PN WO200135255-A2.
XX
PD 17-MAY-2001.

XX 08-NOV-2000; 2000WO-US030753.
PF
XX 09-NOV-1999; 99US-00437738.
PR
XX (UYRQ) UNIV ROCKEFELLER.
PA
XX Sali A, Sanchez R, Melo F;
PI
XX WPI; 2001-335968/35.
DR
XX
PT Three dimensional model of protein is generated by searching database(s)
PT to find proteins similar to selected protein amino acid sequence and
PT selecting best model based on selected proteins.
XX Example 1; Page 51; 124pp; English.
PS
XX The present invention describes a computerised process of generating a
CC three dimensional model of a protein, involving searching databases to
CC find proteins having more than a preset similarity to the protein of
CC interest. The process can be used to model protein structures. The
CC present sequence is a peptide fragment used to demonstrate the invention
XX
SQ Sequence 10 AA;

Query Match 86.6%; Score 58; DB 4; Length 10;
Best Local Similarity 80.0%; Pred.No. 4.3;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
|| |||||:
Db 1 PPAYPPPPVP 10

RESULT 26
ADK70159
ID ADK70159 standard; peptide; 10 AA.
XX
AC ADK70159;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human SH3 binding protein 2 SH3 ligand motif.
XX
KW exon; mutant; osteopathic; cytostatic; gene therapy; SH3-binding protein;
KW SH3BP2; mutation; diagnosis; bone homeostasis; cherubism; bone tumor.
XX
OS Homo sapiens.
XX
PN WO2003025197-A2.
XX
PD 27-MAR-2003.
XX
PF 01-FEB-2002; 2002WO-US019164.
XX
PR 02-FEB-2001; 2001US-0266129P.
XX
PA (HARD) HARVARD COLLEGE.
PA (FORS-) FORSYTH DENTAL INFIRMARY FOR CHILDREN.
XX
PI Tiziani V, Reichenberger E, Ueki Y, Olsen BR;
XX
DR WPI; 2003-371820/35.
XX
PT New mutant SH3BP2 nucleic acid molecule or polypeptide, useful for
PT diagnosing and treating disorders of bone homeostasis, such as cherubism
PT or bone tumor.
XX
PS Example 1; SEQ ID NO 103; 70pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule comprising a
CC mutant SH3-binding protein and its encoding DNA, nucleic acid molecules
CC which hybridize under stringent conditions to it, a nucleic acid

CC molecules that differ from it in codon sequence due to the degeneracy of
CC the genetic code; and complements of these, provided that the nucleic
CC acid molecule is not a human 'WT' SH3BP2 full length sequence. The mutant
CC is especially selected from a genetic mutant domain; 'H' family mutation;
CC 'K' family mutation; 'A,B' family mutation; 'C, F, J, M, O' family
CC mutation; 'L' family mutation; 'G' family mutation; 'N' family mutation
CC or genomic mutant SH3BP2), any of the mutant SH3BP2 exon 9 sequences and
CC a genomic mutant SH3BP2 nucleic acid molecule. The nucleic acid molecule
CC and polypeptide are useful for diagnosing and treating disorders of bone
CC homeostasis, such as cherubism or a bone tumor. This sequence corresponds
CC to the wild type SH3-binding protein ligand motif.
XX
SQ Sequence 10 AA;

Query Match 86.6%; Score 58; DB 7; Length 10;
Best Local Similarity 80.0%; Pred.No. 4.3;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
|| |||||:
Db 1 PPAYPPPPVP 10

RESULT 27
ABG75623
ID ABG75623 standard; peptide; 11 AA.
XX
AC ABG75623;
XX
DT 25-APR-2003 (first entry)
XX
DE Human Abelson protein tyrosine kinase ligand #2.
XX
KW Abelson protein tyrosine kinase; Abl; SH2; SH3; ligand; protein ligation;
KW intein-chitin binding domain; CBD; N-terminal cysteine;
KW unoxidised sulphhydryl side chain; protein-chip; semi-synthetic protein.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 11
FT /label= OTHER
FT /note= "C-terminal amide"
XX
PN US2002151006-A1.
XX
PD 17-OCT-2002.
XX
PF 12-JUL-2001; 2001US-00904117.
XX
PR 13-NOV-1997; 97US-0065391P.
PR 24-JUL-1998; 98US-0093990P.
PR 13-NOV-1998; 98US-00191890.
XX
PA (MUIR/) MUIR T W.
PA (COLE/) COLE P A.
PA (FRIE/) FRIEDMAN J M.
PA (SOND/) SONDHI D.
PA (SEVE/) SEVERINOV K.
XX
PI Muir TW, Cole PA, Friedman JM, Sondhi D, Severinov K;
XX
DR WPI; 2003-238034/23.
XX
PT Cleaving recombinantly expressed proteins bound to intein-chitin binding
PT domains, and ligating to peptides with N-terminal cysteines having
PT unoxidized sulphhydryl side chain, useful for producing semi-synthetic
PT proteins.
XX
PS Example 3; Page 15; 35pp; English.
XX
CC The invention relates to cleaving a recombinantly expressed protein bound
CC to an intein-chitin binding domain (CBD) and ligating the cleaved protein

CC to a peptide with an N-terminal cysteine having an unoxidised sulphhydryl
CC side chain, by contacting the bound protein with a peptide in the
CC presence of conjugated thiol to effect cleavage of the protein from the
CC intein-CBD and production of a C-terminal thioester of the protein which
CC undergoes intramolecular rearrangement to form amide bond linking the
CC protein to the peptide. The method is used for linking recombinantly
CC expressed protein bound to an intein-CBD to a peptide with N-terminal
CC cysteine having an unoxidised sulphhydryl side chain. The method is
CC useful for producing a protein-chip which is useful for diagnostic
CC screening for a specific protein, antibody or antigen and for preparing
CC semi-synthetic proteins of any size. The protein-chip is also useful for
CC identifying presence of protein in a sample. The method is also useful
CC for generating recombinant proteins e.g. cytotoxic recombinant proteins,
CC recombinant proteins partially labeled with a detectable marker, in
CC nuclear magnetic resonance (NMR) spectroscopy, where proteins used in the
CC method are segmentally labeled, and for generating semi-synthetic
CC proteins to facilitate two recombinant, folded proteins to be ligated
CC together. Ligated SH2 and SH3 domains of the human Abelson protein
CC tyrosine kinase (Abl) were produced using the method of the invention.
CC The present sequence is a ligand for Abl used to determine if the ligated
CC domains have folded correctly

XX
SQ Sequence 11 AA;

Query Match 86.6%; Score 58; DB 6; Length 11;
Best Local Similarity 80.0%; Pred. No. 4.7;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
|| |||||:
Db 1 PPAYPPPPVP 10

RESULT 28
AAB12074
ID AAB12074 standard; peptide; 40 AA.
XX
AC AAB12074;
XX
DT 17-JAN-2001 (first entry)
XX
DE SH3 binding domain from 3BP2.
XX
KW Rb-interacting zinc finger; RIZ; retinoblastoma; heart disease;
KW cell proliferation; cell differentiation; tissue repair;
KW transcription regulator; breast cancer; gene therapy; melanoma;
KW neuroblastoma; leukaemia; Parkinson's disease; Huntingdon's disease;
KW Alzheimer's disease; paralysis; motor neurone disorder;
XX SH3 binding domain.
XX
OS Unidentified.
XX
PN US6069231-A.
XX
PD 30-MAY-2000.
XX
PF 18-AUG-1995; 95US-00516859.
XX
PR 18-AUG-1994; 94US-00292683.
PR 06-MAR-1995; 95US-00399411.
XX
PA (LJOL-) LA JOLLA CANCER RES FOUND.
XX
XX Huang S;
PI
XX WPI; 2000-410879/35.
XX
DR New PR domain peptides comprising amino acid sequences from, for example
PT retinoblastoma-interacting zinc finger, or egl-43 proteins, for
PT regulating gene transcription and controlling cell proliferation and
PT differentiation.
XX
PS Disclosure; Fig 2; 91pp; English.

XX
CC The present sequence is the SH3 binding domain from 3BP2 protein. A SH3
CC binding domain is also found in rat RIZ (AAB12028). RIZ is retinoblastoma
CC (Rb)-interacting zinc finger protein. RIZ is a nuclear phosphoprotein
CC that acts as a cell differentiation factor. RIZ can modulate cell growth
CC by binding to Rb protein, which is involved in regulating cell
CC proliferation. In addition, RIZ can act to regulate transcription. RIZ
CC functions to maintain cells in the G1 phase of the cell cycle, by
CC interacting with Rb through the cr2 domain of RIZ. Rat RIZ protein
CC contains a number of GTPase motifs (see AAB12037 to AAB12056 and AAB12099
CC to AAB12104). RIZ protein is a PR domain protein and is present primarily
CC in the cell nucleus. RIZ gene mutations may be implicated in various
CC cancers such as melanoma, neuroblastoma, leukaemia and breast cancer, and
CC so the RIZ gene may be used in gene therapy for these disorders. Since
CC RIZ protein is implicated in cell cycle arrest, inhibition of RIZ
CC activity may be useful in neurodegenerative disorder therapy e.g. for
CC Parkinson's, Huntingdon's or Alzheimer's disease, paralysis or motor
CC neurone disorders, or cardiac disorders e.g. heart disease, where the
CC ability to induce neural/ cardiac tissue proliferation would be useful.
CC The present sequence was used for sequence homology comparison
XX
SQ Sequence 40 AA;

Query Match 86.6%; Score 58; DB 3; Length 40;
Best Local Similarity 80.0%; Pred. No. 13;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
|| |||||:
Db 1 PPAYPPPPVP 10

RESULT 29
ADE84647
ID ADE84647 standard; protein; 40 AA.
XX
AC ADE84647;
XX
DT 29-JAN-2004 (first entry)
XX
DE 3BP2 SH3 binding protein motif.
XX
KW cytostatic; tumor; gene therapy;
KW retinoblastoma protein-interacting zinc protein; RIZ; neuroblastoma;
KW melanoma; cardiac cell growth inducer; neuronal cell growth inducer;
KW SH3 binding protein motif.
XX
OS Unidentified;.
XX
PN US6468985-B1.
XX
PD 22-OCT-2002.
XX
PF 17-MAR-2000; 2000US-00528706.
XX
PR 18-AUG-1994; 94US-00292683.
PR 06-MAR-1995; 95US-00399411.
PR 18-AUG-1995; 95US-00516859.
XX
PA (BURN-) BURNHAM INST.
XX
PI Huang S;
XX
DR WPI; 2003-147106/14.
XX
PT Reducing the growth of tumor cell having mammalian retinoblastoma protein
PT -interacting zinc finger proteins in a subject, comprises administering
PT nucleic acid molecule encoding RIZ at or adjacent to the site of tumor.
XX
PS Disclosure; SEQ ID NO 60; 49pp; English.
XX
CC The invention describes a method of reducing growth of a tumour cell
CC having a mutant mammalian retinoblastoma (Rb) protein-interacting zinc

CC finger (RIZ) protein in a subject The method comprises administering a
CC nucleic acid molecule (I) encoding RIZ at or adjacent to tumour site. (I)
CC is administered at the site of the tumour. The method is useful for
CC reducing the growth of a tumour cell e.g. neuroblastoma or melanoma cell
CC having mutant RIZ protein. The method is useful for inducing growth of a
CC cardiac cell or a neuronal cell in a subject, and for effecting normal
CC growth control to a tumour cell or causing differentiation of tumour
CC cells. This is the amino acid sequence of a SH3 binding protein motif
CC used in a comparison with RIZ protein.
XX
SQ Sequence 40 AA;

Query Match 86.6%; Score 58; DB 7; Length 40;
Best Local Similarity 80.0%; Pred. No. 13;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
|| |||||:|
Db 1 PPAYPPPPVVP 10

RESULT 30
ADQ67211
ID ADQ67211 standard; protein; 467 AA.

XX
AC ADQ67211;
XX
DT 07-OCT-2004 (first entry)
XX
DE Novel human protein sequence #2184.
XX
KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;
KW gene therapy; diagnostic marker; morbid state; osteoporosis;
KW neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
KW cancer.

XX
OS Homo sapiens.
XX
PN EP1440981-A2.
XX
PD 28-JUL-2004.

XX
PF 21-JAN-2004; 2004EP-00001196.
XX
PR 21-JAN-2003; 2003JP-00102206.
PR 09-MAY-2003; 2003JP-00131392.

XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Nagai K, Irie R;

XX
DR WPI; 2004-535376/52.
DR N-PSDB; ADQ65023.

XX
PT Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX
PS Claim 1; SEQ ID NO 4372; 2449pp; English.

XX
CC The invention relates to 2495 novel polynucleotides (I) and their encoded
CC polypeptides, sequences hybridizing to these nucleotides, sequences
CC encoding partial polypeptides and sequences having 70% or 90% identity to
CC the nucleotide and protein sequences. The nucleotides and polypeptides
CC are useful as diagnostic markers or therapeutic target for the diseases
CC or morbid states. They are also useful for treating osteoporosis,
CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,
CC dementia and various cancers. This sequence corresponds to a protein
CC sequence of the invention.

XX
SQ Sequence 467 AA;
Query Match 86.6%; Score 58; DB 8; Length 467;

Best Local Similarity 80.0%; Pred. No. 1e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 PPPYPPPPPIP 10
|| |||||:|
Db 176 PPAYPPPPVVP 185

RESULT 31
ADQ96504
ID ADQ96504 standard; protein; 492 AA.

XX
AC ADQ96504;
XX
DT 07-OCT-2004 (first entry)
XX
DE T cell activation associated protein #341.
XX
KW antiallergic; antiarthritic; antiasthmatic; antidiabetic; anti-HIV;
KW antimicrobial; antirheumatic; immunosuppressive; neuroprotective;
KW gene therapy; T cell activation; diagnosis; autoimmune disease;
KW rheumatoid arthritis; asthma; multiple sclerosis; diabetes;
KW allergic disease; infectious disease; AIDS; chronic rejection; organ;
KW bone-marrow transplant.

XX
OS Homo sapiens.
XX
PN WO2004058805-A2.

XX
PD 15-JUL-2004.
XX
PF 25-DEC-2003; 2003WO-JP016715.

XX
PR 26-DEC-2002; 2002JP-00376365.
PR 27-DEC-2002; 2002US-0436473P.
PR 25-APR-2003; 2003JP-00122113.
PR 28-APR-2003; 2003US-0465792P.
PR 21-OCT-2003; 2003JP-00360559.
PR 22-OCT-2003; 2003US-0512846P.

XX
PA (ASAH-) ASahi Kasei Pharma Corp.

XX
PI Matsuda A, Yoneta S;

XX
DR WPI; 2004-593134/57.
DR N-PSDB; ADQ96503.

XX
PT New purified protein involved in T cell activation, useful for
PT diagnosing, preventing and/or treating acquired immunodeficiency
PT syndrome, autoimmune (e.g. rheumatoid arthritis, and diabetes), allergic
PT and infectious diseases.

XX
PS Claim 1; SEQ ID NO 682; 2828pp; English.

XX
CC The invention relates to purified proteins and genes encoding them, that
CC are involved in T cell activation (I) and has an amino acid deletion,
CC substitution or addition in the amino acid sequences. The methods and
CC compositions of the present invention are useful for the diagnosis,
CC prevention and/or treatment of autoimmune disease (rheumatoid arthritis,
CC asthma, multiple sclerosis and diabetes), allergic disease, infectious
CC disease, AIDS, and acute or chronic rejection at organ transplant or bone
CC -marrow transplant. This sequence corresponds to a protein involved in T
CC cell activation.

XX
SQ Sequence 492 AA;

Query Match 86.6%; Score 58; DB 8; Length 492;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
|| |||||:|
Db 201 PPAYPPPPVVP 210

RESULT 32
ADQ66376
ID ADQ66376 standard; protein; 536 AA.
XX AC ADQ66376;
XX DT 07-OCT-2004 (first entry)
XX DE Novel human protein sequence #1349.
XX KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;
KW gene therapy; diagnostic marker; morbid state; osteoporosis;
KW neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
KW cancer.
XX OS Homo sapiens.
XX PN EP1440981-A2.
XX PD 28-JUL-2004.
XX PF 21-JAN-2004; 2004EP-00001196.
XX PR 21-JAN-2003; 2003JP-00102206.
XX PR 09-MAY-2003; 2003JP-00131392.
XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Nagai K, Irie R;
XX WPI; 2004-535376/52.
DR N-PSDB; ADQ64188.
XX Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX Claim 1; SEQ ID NO 3537; 2449pp; English.
XX The invention relates to 2495 novel polynucleotides (I) and their encoded
CC polypeptides, sequences hybridizing to these nucleotides, sequences
CC encoding partial polypeptides and sequences having 70% or 90% identity to
CC the nucleotide and protein sequences. The nucleotides and polypeptides
CC are useful as diagnostic markers or therapeutic target for the diseases
CC or morbid states. They are also useful for treating osteoporosis,
CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,
CC dementia and various cancers. This sequence corresponds to a protein
CC sequence of the invention.
XX Sequence 536 AA;
Query Match 86.6%; Score 58; DB 8; Length 536;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 PPPYPPPPPIP 10
Db 176 PPAYPPPPVP 185
RESULT 33
ADC71301
ID ADC71301 standard; protein; 561 AA.
XX AC ADC71301;
XX DT 18-DEC-2003 (first entry)
XX DE Human colon specific protein sequence DEX0235_129 (SeqID 129).
XX KW human; gene; ss; neoplastic colorectal; colon cancer;

KW non-cancerous disease; gene therapy; transgenic; DEX0235_129.
XX OS Homo sapiens.
XX PN WO2003020934-A1.
XX PD 13-MAR-2003.
XX PF 29-AUG-2002; 2002WO-US027737.
XX PR 31-AUG-2001; 2001US-0316258P.
XX (DIAD-) DIADEXUS INC.
XX PI Sun Y, Liu C, Ghosh MG;
XX WPI; 2003-300891/29.
DR N-PSDB; ADC71368.
XX Novel colon specific polypeptides and nucleic acids, useful for
PT identifying, diagnosing, monitoring, staging, imaging and treating colon
PT cancer and non-cancerous disease states in colon tissue.
XX Claim 11; SEQ ID NO 129; 262pp; English.
XX This invention relates to novel nucleic acid molecules and the encoded
CC polypeptides, which are present in normal and neoplastic colorectal
CC cells. Specifically, it refers to antibodies of these colon specific
CC polypeptides, as well as antagonists and agonists thereof that can be
CC used to treat colon cancer and also non-cancerous diseases states of the
CC colon. The present invention describes methods useful for the diagnosis
CC and monitoring of colon cancer metastases in a patient, by determining
CC the concentration of these colon specific proteins in a patient sample.
CC Furthermore, they are also used for gene therapy purposes, the production
CC of transgenic animals and cells, as well as producing engineered colon
CC tissue for treatment and research. This polypeptide sequence is a human
CC colon specific protein sequence of the invention.
XX Sequence 561 AA;
Query Match 86.6%; Score 58; DB 7; Length 561;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 PPPYPPPPPIP 10
Db 201 PPAYPPPPVP 210
RESULT 34
ADK70062
ID ADK70062 standard; protein; 561 AA.
XX AC ADK70062;
XX DT 06-MAY-2004 (first entry)
XX DE Mutant human SH3 binding protein 2.
XX KW osteopathic; cytostatic; gene therapy; SH3-binding protein; SH3BP2;
KW mutation; diagnosis; bone homeostasis; cherubism; bone tumor; mutant;
KW mutein.
XX OS Homo sapiens.
XX OS Synthetic.
XX PN WO2003025197-A2.
XX PD 27-MAR-2003.
XX PF 01-FEB-2002; 2002WO-US019164.
XX PR 02-FEB-2001; 2001US-02666129P.

XX (HARD) HARVARD COLLEGE.
PA (FORS-) FORSYTH DENTAL INFIRMARY FOR CHILDREN.
XX
PI Tiziani V, Reichenberger E, Ueki Y, Olsen BR;
XX
DR WPI; 2003-371820/35.
DR N-PSDB; ADK70061.
XX
PT New mutant SH3BP2 nucleic acid molecule or polypeptide, useful for
PT diagnosing and treating disorders of bone homeostasis, such as cherubism
PT or bone tumor.
XX
PS Claim 9; SEQ ID NO 6; 70pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule comprising a
CC mutant SH3-binding protein and its encoding DNA, nucleic acid molecules
CC which hybridize under stringent conditions to it, a nucleic acid
CC molecules that differ from it in codon sequence due to the degeneracy of
CC the genetic code; and complements of these, provided that the nucleic
CC acid molecule is not a human 'WT' SH3BP2 full length sequence. The mutant
CC is especially selected from a genetic mutant domain; 'H' family mutation;
CC 'K' family mutation; 'A,B' family mutation; 'C, F, J, M, O' family
CC mutation; 'L' family mutation; 'G' family mutation; 'N' family mutation
CC or genomic mutant SH3BP2), any of the mutant SH3BP2 exon 9 sequences and
CC a genomic mutant SH3BP2 nucleic acid molecule. The nucleic acid molecule
CC and polypeptide are useful for diagnosing and treating disorders of bone
CC homeostasis, such as cherubism or a bone tumor. This sequence corresponds
CC to the mutant full length SH3-binding protein.
XX
SQ Sequence 561 AA;

Query Match 86.6%; Score 58; DB 7; Length 561;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
||| |||||:|
Db 201 PPAYPPPPVP 210

RESULT 35
ADK70058
ID ADK70058 standard; protein; 561 AA.
XX
AC ADK70058;
XX
DT 06-MAY-2004 (first entry)
XX
DE Wild type human SH3 binding protein 2.
XX
KW osteopathic; cytostatic; gene therapy; SH3-binding protein; SH3BP2;
KW mutation; diagnosis; bone homeostasis; cherubism; bone tumor.
XX
OS Homo sapiens.
XX
PN WO2003025197-A2.
XX
PD 27-MAR-2003.
XX
PF 01-FEB-2002; 2002WO-US019164.
XX
PR 02-FEB-2001; 2001US-0266129P.
XX

(HARD) HARVARD COLLEGE.
(FORS-) FORSYTH DENTAL INFIRMARY FOR CHILDREN.
XX
PI Tiziani V, Reichenberger E, Ueki Y, Olsen BR;
XX
DR WPI; 2003-371820/35.
DR N-PSDB; ADK70057.

PT New mutant SH3BP2 nucleic acid molecule or polypeptide, useful for

PT diagnosing and treating disorders of bone homeostasis, such as cherubism
PT or bone tumor.
XX
PS Disclosure; SEQ ID NO 2; 70pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule comprising a
CC mutant SH3-binding protein and its encoding DNA, nucleic acid molecules
CC which hybridize under stringent conditions to it, a nucleic acid
CC molecules that differ from it in codon sequence due to the degeneracy of
CC the genetic code; and complements of these, provided that the nucleic
CC acid molecule is not a human 'WT' SH3BP2 full length sequence. The mutant
CC is especially selected from a genetic mutant domain; 'H' family mutation;
CC 'K' family mutation; 'A,B' family mutation; 'C, F, J, M, O' family
CC mutation; 'L' family mutation; 'G' family mutation; 'N' family mutation
CC or genomic mutant SH3BP2), any of the mutant SH3BP2 exon 9 sequences and
CC a genomic mutant SH3BP2 nucleic acid molecule. The nucleic acid molecule
CC and polypeptide are useful for diagnosing and treating disorders of bone
CC homeostasis, such as cherubism or a bone tumor. This sequence corresponds
CC to the wild type full length SH3-binding protein.
XX
SQ Sequence 561 AA;

Query Match 86.6%; Score 58; DB 7; Length 561;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
||| |||||:|
Db 201 PPAYPPPPVP 210

RESULT 36
ADQ96506
ID ADQ96506 standard; protein; 561 AA.
XX
AC ADQ96506;

XX
DT 07-OCT-2004 (first entry)
XX
DE T cell activation associated protein #342.
XX

KW antiallergic; antiarthritic; antiasthmatic; antidiabetic; anti-HIV;
KW antimicrobial; antirheumatic; immunosuppressive; neuroprotective;
KW gene therapy; T cell activation; diagnosis; autoimmune disease;
KW rheumatoid arthritis; asthma; multiple sclerosis; diabetes;
KW allergic disease; infectious disease; AIDS; chronic rejection; organ;
KW bone-marrow transplant.

XX
OS Homo sapiens.
XX
PN WO2004058805-A2.

XX
PD 15-JUL-2004.
XX
PF 25-DEC-2003; 2003WO-JP016715.
XX
PR 26-DEC-2002; 2002JP-00376365.
PR 27-DEC-2002; 2002US-0436473P.
PR 25-APR-2003; 2003JP-00122113.
PR 28-APR-2003; 2003US-0465792P.
PR 21-OCT-2003; 2003JP-00360559.
PR 22-OCT-2003; 2003US-0512846P.

XX (ASAH-) ASahi Kasei PHARMA CORP.
XX
PI Matsuda A, Yoneta S;
XX
DR WPI; 2004-593134/57.
DR N-PSDB; ADQ96505.

XX
PT New purified protein involved in T cell activation, useful for
PT diagnosing, preventing and/or treating acquired immunodeficiency
PT syndrome, autoimmune (e.g. rheumatoid arthritis, and diabetes), allergic

PT and infectious diseases.
XX
PS Claim 1; SEQ ID NO 684; 2828pp; English.
XX
CC The invention relates to purified proteins and genes encoding them, that
CC are involved in T cell activation (I) and has an amino acid deletion,
CC substitution or addition in the amino acid sequences. The methods and
CC compositions of the present invention are useful for the diagnosis,
CC prevention and/or treatment of autoimmune disease (rheumatoid arthritis,
CC asthma, multiple sclerosis and diabetes), allergic disease, infectious
CC disease, AIDS, and acute or chronic rejection at organ transplant or bone
CC -marrow transplant. This sequence corresponds to a protein involved in T
CC cell activation.
XX
SQ Sequence 561 AA;

Query Match 86.6%; Score 58; DB 8; Length 561;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 201 PPAYPPPPVP 210
|| |||||:|

RESULT 37
ABG19592
ID ABG19592 standard; protein; 50 AA.
XX
AC ABG19592;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #19583.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS83779.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 49951; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging

CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 50 AA;

Query Match 85.1%; Score 57; DB 4; Length 50;
Best Local Similarity 80.0%; Pred. No. 21;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 4 PPPLPPPPPLP 13
||| ||||:|

RESULT 38
ABB71558
ID ABB71558 standard; protein; 94 AA.
XX
AC ABB71558;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 41466.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL15661.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Disclosure; SEQ ID NO 41466; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 94 AA;


```
Query Match      85.1%; Score 57; DB 4; Length 94;
Best Local Similarity 80.0%; Pred. No. 35;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
Db 78 PPPLPPPPPLP 87

RESULT 39
AAO02495
ID AAO02495 standard; protein; 98 AA.
XX
AC AAO02495;
XX
DT 06-NOV-2001 (first entry)
DE Human polypeptide SEQ ID NO 16387.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US004927.
XX
PR 28-FEB-2000; 2000US-00515126.
PR 18-MAY-2000; 2000US-00577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-514838/56.
DR N-PSDB; AAI82426.
XX
PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
XX
PS Claim 20; SEQ ID NO 16387; 1399pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO0010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 98 AA;

Query Match      85.1%; Score 57; DB 4; Length 98;
Best Local Similarity 80.0%; Pred. No. 36;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
Db 20 PPPLPPPPPLP 29

RESULT 40
```

```
ABG16489
ID ABG16489 standard; protein; 118 AA.
XX
AC ABG16489;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #16480.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS80676.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 46848; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 118 AA;

Query Match      85.1%; Score 57; DB 4; Length 118;
Best Local Similarity 80.0%; Pred. No. 42;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
Db 11 PPPLPPPPPLP 20

RESULT 41
ABO76127
ID ABO76127 standard; protein; 142 AA.
```

XX ABO76127;
AC
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polypeptide #8302.
XX
KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
PN US6551795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
DR WPI; 2003-615309/58.
DR N-PSDB; ABD09698.
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 24873; 455pp; English.
XX
CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide, of
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 142 AA;

Query Match 85.1%; Score 57; DB 7; Length 142;
Best Local Similarity 80.0%; Pred. No. 48;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
Db 89 PPPSPPPPPVP 98
||| ||||:|

RESULT 42
AAB80893
ID AAB80893 standard; protein; 182 AA.
XX
AC AAB80893;
XX
DT 30-MAY-2001 (first entry)
XX
DE Human ATFX leucine zipper domain.
XX
KW Human; ATFX; gamma aminobutyric acid B receptor; GABA B receptor;
KW transcription factor; leucine zipper.

XX Homo sapiens.
OS
XX
PN WO200116596-A2.
XX
PD 08-MAR-2001.
XX
PF 31-AUG-2000; 2000WO-GB003361.
XX
PR 31-AUG-1999; 99GB-00020569.
PR 12-JAN-2000; 2000GB-00000516.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI White J, Wise A, Marshall F;
XX
DR WPI; 2001-265904/27.
DR N-PSDB; AAF77878.
XX
PT Identifying modulators of gamma aminobutyric acid(GABA)-B receptor
PT mediated activity by monitoring the interaction between GABAB receptor
PT and the CREB/ATF transcription factors in the presence of a test
PT substance.
XX
PS Claim 14; Page 64-65; 71pp; English.
XX
CC The present invention relates to a method for identifying a modulator of
CC gamma aminobutyric acid B (GABA_B) receptor-mediated activity, by
CC monitoring the interaction between a CREB/ATF transcription factor
CC capable of binding to GABA_B receptor. The present sequence is the
CC leucine zipper domain of human ATFX (see AAB80892), which was used in the
CC method of the present invention. ATFX is a member of the CREB/ATF family
CC of transcription factors. ATFX contains a bZIP domain, by which it
CC interacts with the coiled coil domain of GABA_B receptor. Modulators of
CC GABA_B receptor activity are useful for treating central nervous system
CC or peripheral nervous system disorders
XX
SQ Sequence 182 AA;

Query Match 85.1%; Score 57; DB 4; Length 182;
Best Local Similarity 80.0%; Pred. No. 59;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
Db 27 PPPLPPPPPLP 36
||| ||||:|

RESULT 43
AAB80892
ID AAB80892 standard; protein; 216 AA.
XX
AC AAB80892;
XX
DT 30-MAY-2001 (first entry)
XX
DE Human ATFX.
XX
KW Human; ATFX; gamma aminobutyric acid B receptor; GABA B receptor;
KW transcription factor.
XX
OS Homo sapiens.
XX
PN WO200116596-A2.
XX
PD 08-MAR-2001.
XX
PF 31-AUG-2000; 2000WO-GB003361.
XX
PR 31-AUG-1999; 99GB-00020569.
PR 12-JAN-2000; 2000GB-00000516.
XX
PA (GLAX) GLAXO GROUP LTD.

XX White J, Wise A, Marshall F;
PI WPI; 2001-265904/27.
XX N-PSDB; AAF77877.
PT Identifying modulators of gamma aminobutyric acid(GABA)-B receptor
PT mediated activity by monitoring the interaction between GABAB receptor
PT and the CREB/ATF transcription factors in the presence of a test
PT substance.
XX Claim 14; Fig 1; 7lpp; English.
PS The present invention relates to a method for identifying a modulator of
CC gamma aminobutyric acid B (GABA_B) receptor-mediated activity, by
CC monitoring the interaction between a CREB/ATF transcription factor
CC capable of binding to GABA_B receptor. The present invention is human
CC ATFx, which was used in the method of the present invention. ATFx is a
CC member of the CREB/ATF family of transcription factors. ATFx contains a
CC bZIP domain, by which it interacts with the coiled coil domain of GABA_B
CC receptor. Modulators of GABA_B receptor activity are useful for treating
CC central nervous system or peripheral nervous system disorders
XX
SQ Sequence 216 AA;

Query Match 85.1%; Score 57; DB 4; Length 216;
Best Local Similarity 80.0%; Pred. No. 68;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 61 PPPLPPPPPLP 70

RESULT 44
ABB50151
ID ABB50151 standard; protein; 216 AA.
XX
AC ABB50151;
XX
DT 05-FEB-2002 (first entry)
XX
DE Human transcription factor TRFX-2.
XX
KW Human; transcription factor; TRFX; cell proliferative disease;
KW autoimmune disease; inflammation; neurological disease;
KW developmental disorder; cancer; AIDS; infection; cytostatic; anti-HIV;
KW neuroprotective; antiinflammatory; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200172777-A2.
XX
PD 04-OCT-2001.
XX
PF 13-MAR-2001; 2001WO-US008117.
XX
PR 13-MAR-2000; 2000US-0188986P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Hillman JL, Baughn MR, Yue H, Lal P, Lu DAM, Patterson C;
PI Azimzai Y, Bandman O, Tang YT, Mathur P, Shah P, Au-Young J;
PI Reddy R;
XX
DR WPI; 2001-570896/64.
DR N-PSDB; ABA82975.
XX
PT Novel transcription factor polypeptides, used to treat diseases
PT associated with altered activity and expression of TRFX, and to screen
PT for agents capable of modulating its activity.
XX
PS Claim 1; Page 140-141; 327pp; English.

XX The present sequence is the protein sequence for a human transcription
CC factor. The transcription factor and its coding sequence are useful in
CC the diagnosis, treatment and prevention of diseases associated with
CC altered expression of the transcription factor e.g. cell proliferative,
CC autoimmune/inflammatory, neurological and developmental disorders. A
CC number of specific disorders/diseases are given in the specification,
CC including: arteriosclerosis, cirrhosis, hepatitis, cancers, AIDS,
CC allergies, anaemia, asthma, autoimmune thyroiditis, bronchitis, atopic
CC dermatitis, diabetes mellitus, emphysema, Goodpasture's syndrome, gout,
CC Grave's disease, multiple sclerosis, osteoarthritis, pancreatitis,
CC psoriasis, rheumatoid arthritis, systemic lupus erythematosus, ulcerative
CC colitis, uveitis, Alzheimer's disease, Huntington's disease, Parkinson's
CC disease, stroke, and viral, bacterial, fungal and protozoal infections
XX
SQ Sequence 216 AA;

Query Match 85.1%; Score 57; DB 4; Length 216;
Best Local Similarity 80.0%; Pred. No. 68;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 61 PPPLPPPPPLP 70

RESULT 45
ABG19594
ID ABG19594 standard; protein; 227 AA.
XX
AC ABG19594;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #19585.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS83781.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 49953; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food

CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activities. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AEG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 227 AA;

Query Match 85.1%; Score 57; DB 4; Length 227;
Best Local Similarity 80.0%; Pred. No. 71;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
||| ||||:|
Db 64 PPPLPPPLP 73

RESULT 46
AAB80896
ID AAB80896 standard; protein; 282 AA.

XX
AC AAB80896;
XX
DT 30-MAY-2001 (first entry)
XX
DE Human ATFX with N-terminal extension.

XX
KW Human; ATFX; gamma aminobutyric acid B receptor; GABA B receptor;
KW transcription factor.

XX
OS Homo sapiens.

XX
PN WO200116596-A2.

XX
PD 08-MAR-2001.

XX
PF 31-AUG-2000; 2000WO-GB003361.

XX
PR 31-AUG-1999; 99GB-00020569.

XX
PR 12-JAN-2000; 2000GB-00000516.

XX
PA (GLAX) GLAXO GROUP LTD.

XX
PI White J, Wise A, Marshall F;

XX
DR WPI; 2001-265904/27.

XX
DR N-PSDB; AAF77880.

XX
PT Identifying modulators of gamma aminobutyric acid(GABA)-B receptor
PT mediated activity by monitoring the interaction between GABAB receptor
PT and the CREB/ATF transcription factors in the presence of a test
PT substance.

XX
PS Claim 14; Fig 9; 71pp; English.

XX
CC The present invention relates to a method for identifying a modulator of
CC gamma aminobutyric acid_B (GABA_B) receptor-mediated activity, by
CC monitoring the interaction between a CREB/ATF transcription factor
CC capable of binding to GABA_B receptor. The present invention is human
CC ATFX, which was used in the method of the present invention. ATFX is a
CC member of the CREB/ATF family of transcription factors. ATFX contains a
CC bZIP domain, by which it interacts with the coiled coil domain of GABA_B
CC receptor. Modulators of GABA_B receptor activity are useful for treating
CC central nervous system or peripheral nervous system disorders. The
CC present sequence has an N-terminal extension compared to AAB80892

XX

SQ Sequence 282 AA;

Query Match 85.1%; Score 57; DB 4; Length 282;
Best Local Similarity 80.0%; Pred. No. 85;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
||| ||||:|
Db 127 PPPLPPPLP 136

RESULT 47
AAB95859
ID AAB95859 standard; protein; 282 AA.

XX
AC AAB95859;

XX
DT 26-JUN-2001 (first entry)

XX
DE Human protein sequence SEQ ID NO:18922.

XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX
OS Homo sapiens.

XX
PN EP1074617-A2.

XX
PD 07-FEB-2001.

XX
PF 28-JUL-2000; 2000EP-00116126.

XX
PR 29-JUL-1999; 99JP-00248036.

XX
PR 27-AUG-1999; 99JP-00300253.

XX
PR 11-JAN-2000; 2000JP-00118776.

XX
PR 02-MAY-2000; 2000JP-00183767.

XX
PR 09-JUN-2000; 2000JP-00241899.

XX
PA (HELI-) HELIX RES INST.

XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX
DR WPI; 2001-318749/34.

XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

XX
PS Claim 8; SEQ ID NO 18922; 2537pp + Sequence Listing; English.

XX
CC The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention

XX SQ Sequence 282 AA; Query Match 85.1%; Score 57; DB 4; Length 282; Best Local Similarity 80.0%; Pred. No. 85; Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
| | | | | | | | | |
Db 127 PPPLPPPPPLP 136

RESULT 48
ABB81169
ID ABB81169 standard; protein; 282 AA.
XX
AC ABB81169;
XX
DT 25-NOV-2002 (first entry)
XX
DE Human activating transcription factor 5 (ATF5) sequence.
XX
KW Epstein-Barr virus; EBV; TRRAP; cofactor; ATF5; elf1; SUPT6H; human;
KW activating transcription factor 5; translation initiation factor;
KW structure regulator; gene therapy; virucide; cytostatic.
XX
OS Homo sapiens.
XX
PN WO200259606-A2.
XX
PD 01-AUG-2002.
XX
PF 23-JAN-2002; 2002WO-BF000659.
XX
PR 23-JAN-2001; 2001EP-00101452.
XX
PA (GPCB-) GPC BIOTECH AG.
PA (GSFU-) GSF FORSCHUNGSZENTRUM UMWELT & GESUNDHEI.
XX
PI Kempkes B, Ivanov I, Machl A, Bornkamm G, Santak M;
XX
WPI; 2002-643349/69.
DR N-PSDB; ABQ79514.
XX
Identifying antagonists or inhibitors for treating or preventing Epstein-Barr virus infections, comprises testing for inhibition or reduction of transcription of a specific gene, and of translation of mRNA transcribed from the gene.
XX
PS Disclosure; Page 34; 45pp; English.
XX
The invention relates to identifying antagonists or inhibitors of EBV (Epstein-Barr virus) infection. The method involves testing candidate antagonists or inhibitors for the inhibition or reduction of transcription of a gene (I), and of translation of mRNA transcribed from (I), and determining whether the antagonists or inhibitors, or samples containing them test positive. (I) encodes a polypeptide selected from TRRAP (cofactor), activating transcription factor 5 (ATF5), translation initiation factor (elf1), and structure regulator (SUPT6H), their fragments, derivatives or orthologues. The method is used for identifying antagonists or inhibitors which are useful in treating or preventing EBV infections or other related diseases such as EBV-associated haemophagocytic syndrome, chronic active EBV infection, T-cell lymphoma, natural killer cell leukemia/lymphoma, lymphoproliferative diseases in immunocompromised hosts, Hodgkin's disease, pyothorax-associated B-cell lymphoma, smooth muscle tumours, gastric carcinoma or Burkitt's lymphoma. They are also useful in the preparation of pharmaceutical composition for treating or preventing the EBV infection or related diseases. The polypeptides are useful for identifying the antagonists or inhibitors of EBV infection. Conditional mutants in the genes encoding the polypeptides are useful as markers of potentially EBV infected cells. The present sequence represents the human ATF5 amino acid sequence (GenBank accession no. NM_012068)

XX SQ Sequence 282 AA; Query Match 85.1%; Score 57; DB 5; Length 282; Best Local Similarity 80.0%; Pred. No. 85; Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
| | | | | | | | | |
Db 127 PPPLPPPPPLP 136

RESULT 49
ADE60923
ID ADE60923 standard; protein; 282 AA.
XX
AC ADE60923;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human Protein Q9Y2D1, SEQ ID NO 6837.
XX
KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
WPI; 2003-268312/26.
XX
New composition comprising two or more isolated polypeptides, useful for preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.
XX
The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene therapy). The sequence presented is a human protein (shown in Table 2 of

CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX

SQ Sequence 282 AA;

Query Match 85.1%; Score 57; DB 7; Length 282;
Best Local Similarity 80.0%; Pred. No. 85;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPIP 10
||| |||:|
Db 127 PPPLPPPLP 136

RESULT 50
ADP12591
ID ADP12591 standard; protein; 282 AA.
XX
AC ADP12591;
XX
DT 12-AUG-2004 (first entry)
XX
DE Protein encoded by mRNA of the invention #201.
XX
KW transplant rejection; immune system; rheumatoid arthritis; lupus;
KW inflammatory bowel disease; multiple sclerosis; HIV; AIDS.
XX
OS Homo sapiens.
XX
PN WO2004042346-A2.
XX
PD 21-MAY-2004.
XX
PF 24-APR-2003; 2003WO-US012946.
XX
PR 24-APR-2002; 2002US-00131831.
PR 20-DEC-2002; 2002US-00325899.
XX

PA (EXPR-) EXPRESSION DIAGNOSTICS INC.
XX
PI Wohlgemuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M;
PI Rosenberg S;
XX
DR WPI; 2004-400724/37.
XX
PT Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,
PT pancreas, pancreatic islet, lung, bone marrow or stem cell transplant
PT rejection, in an individual, comprises detecting the expression level of
PT the genes.
XX

PS Claim 65; SEQ ID NO 2600; 1762pp; English.
XX
CC The present invention relates to diagnosing or monitoring transplant
CC rejection, e.g. cardiac or kidney transplant rejection, in an individual
CC comprises detecting the expression level of one or more genes. The
CC methods, system and kits are useful in diagnosing or monitoring
CC transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic
CC islet, lung, bone marrow or stem cell transplant rejection,
CC xenotransplant rejection or mechanical organ replacement rejection, in an
CC individual. The method is also useful in assessing the immune status of
CC an individual. The methods are also useful in diagnosing and monitoring
CC diseases that involve the immune system, e.g. rheumatoid arthritis,
CC lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or
CC viral, bacterial or fungal infection. The present sequence represents a
CC protein that is encoded by the mRNA of the invention.
XX

SQ Sequence 282 AA;

Query Match 85.1%; Score 57; DB 8; Length 282;
Best Local Similarity 80.0%; Pred. No. 85;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPIP 10
||| |||:|
Db 127 PPPLPPPLP 136

Search completed: April 6, 2006, 09:31:16
Job time : 168.211 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2006, 09:35:13 ; Search time 21.7895 Seconds
(without alignments)
52.989 Million cell updates/sec

Title: US-10-632-388-312
Perfect score: 67
Sequence: 1 PPPYPPPIPPX 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : PIR 80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58	86.6	559	2 I49444	SH3 binding protei
2	57	85.1	172	2 D41132	collagen-related p
3	57	85.1	389	2 S27200	proline-rich prote
4	57	85.1	820	2 T00645	hypothetical prote
5	56	83.6	222	2 T43500	hypothetical prote
6	56	83.6	393	2 JC5614	RNB6 protein - rat
7	56	83.6	563	2 A87378	hypothetical prote
8	56	83.6	744	1 QQBEA7	UL69 protein - hum
9	56	83.6	1733	1 B45344	probable nuclear a
10	56	83.6	1953	2 S63244	BN11 protein - yea
11	56	83.6	1958	2 B40505	hypothetical prote
12	55	82.1	182	2 T30760	hypothetical prote
13	55	82.1	211	2 S28304	hypothetical prote
14	55	82.1	281	2 I38707	Fas ligand - human
15	55	82.1	314	2 T48514	hypothetical prote
16	55	82.1	367	1 S02193	cellular tumor ant
17	55	82.1	433	2 T07910	hydroxyproline-ric
18	55	82.1	442	2 T34018	hypothetical prote
19	55	82.1	473	2 B85187	glycoprotein homol
20	55	82.1	520	2 E97813	WASP, N-WASP, MENA
21	55	82.1	529	2 I38607	p53-binding protei
22	55	82.1	662	2 D40228	neurexin II-beta p
23	55	82.1	681	2 JC5929	serine/arginine-ri
24	55	82.1	691	2 T46476	hypothetical prote
25	55	82.1	798	2 D96563	probable bZIP prot
26	55	82.1	907	2 E96636	hypothetical prote
27	55	82.1	994	2 S19595	chloride channel p
28	55	82.1	1058	2 T13286	cappuccino Gene pr
29	55	82.1	1064	2 T13963	formin related pro

30	55	82.1	1171	2 T17454	diaphanous-related
31	55	82.1	1201	2 G86441	unknown protein [i
32	55	82.1	1206	2 S24407	formin isoform IV
33	55	82.1	1255	2 T31065	diaphanous protein
34	55	82.1	1375	2 S48375	hypothetical prote
35	55	82.1	1468	2 S11515	formin - mouse
36	55	82.1	1795	2 F97713	190K antigen precu
37	55	82.1	2783	1 A41948	alpha-fetoprotein
38	54	80.6	406	2 T28957	hypothetical prote
39	54	80.6	710	2 T31502	hypothetical prote
40	54	80.6	1091	2 T13170	diaphanous protein
41	53	79.1	160	2 D48232	cysteine-rich exte
42	53	79.1	161	2 E48232	cysteine-rich exte
43	53	79.1	165	2 C48232	cysteine-rich exte
44	53	79.1	196	2 B48232	cysteine-rich exte
45	53	79.1	209	2 A48232	cysteine-rich exte
46	53	79.1	744	2 E86255	hypothetical prote
47	53	79.1	1127	2 T32404	hypothetical prote
48	53	79.1	1258	2 JC5765	inositol polyphosp
49	53	79.1	1316	2 T00381	KIAA0633 protein -
50	52	77.6	122	2 PQ0452	extensin-like prot
51	52	77.6	138	2 C96734	hypothetical prote
52	52	77.6	144	2 E86364	hypothetical prote
53	52	77.6	149	2 A41132	collagen-related p
54	52	77.6	151	2 S10084	hypothetical 16K p
55	52	77.6	151	2 JQ1686	extensin-like prot
56	52	77.6	153	2- S17837	sorbin - pig
57	52	77.6	186	2 C41132	collagen-related p
58	52	77.6	257	2 S32101	PHLP5A protein - c
59	52	77.6	267	2 T15645	hypothetical prote
60	52	77.6	269	2 C84707	hypothetical prote
61	52	77.6	302	2 A96661	unknown protein, 8
62	52	77.6	359	2 T13478	hypothetical prote
63	52	77.6	378	2 T21651	hypothetical prote
64	52	77.6	383	2 T39597	probable inositol
65	52	77.6	421	1 S11674	acrosin (EC 3.4.21
66	52	77.6	428	2 E71415	probable coll wall
67	52	77.6	464	2 S22697	extensin - Volvox
68	52	77.6	487	2 S42442	nuclear protein EB
69	52	77.6	494	2 B96534	hypothetical prote
70	52	77.6	543	2 S25128	61K protein - Auto
71	52	77.6	588	2 T45564	hypothetical prote
72	52	77.6	645	2 A71416	hypothetical prote
73	52	77.6	649	2 JN0809	drebrin E (clone g
74	52	77.6	651	2 T14763	hypothetical prote
75	52	77.6	666	2 B70803	hypothetical prote
76	52	77.6	684	2 T36771	probable integral
77	52	77.6	757	2 B75437	ABC transporter, A
78	52	77.6	768	2 A87722	protein Zc123.1 [i
79	52	77.6	856	2 D86258	protein F5011.10 [
80	52	77.6	941	2 A86404	probable protein A
81	52	77.6	1013	2 T46422	hypothetical prote
82	52	77.6	1021	2 T23252	hypothetical prote
83	52	77.6	1132	2 A35098	MHC class III hist
84	52	77.6	1388	2 T00063	hypothetical prote
85	52	77.6	1585	2 T31611	hypothetical prote
86	52	77.6	2123	2 F86348	hypothetical prote
87	52	77.6	2453	2 S60254	nuclear receptor c
88	52	77.6	2706	2 T28155	variant-specific s
89	51	76.1	70	2 A96650	protein F2401.6 [i
90	51	76.1	134	2 D84672	hypothetical prote
91	51	76.1	185	2 D75602	hypothetical prote
92	51	76.1	199	2 S14981	extensin class I (
93	51	76.1	221	2 T07176	extensin homolog -
94	51	76.1	250	2 T16342	hypothetical prote
95	51	76.1	277	2 H87552	hypothetical prote
96	51	76.1	278	2 A49266	fas ligand - rat
97	51	76.1	289	2 A43562	homeotic protein H
98	51	76.1	304	2 H86332	T20H2.26 protein -
99	51	76.1	306	2 T06435	ribonuclease S5 ho
100	51	76.1	318	2 T29479	hypothetical prote
101	51	76.1	348	2 AB3260	hypothetical membr
102	51	76.1	391	2 S69192	serine O-acetyltra

103 51 76.1 393 2 PQ0479 pistil extensin-li
104 51 76.1 413 2 H87604 hypothetical prote
105 51 76.1 438 2 T12494 hypothetical prote
106 51 76.1 446 2 T07907 hydroxyproline-ric
107 51 76.1 447 2 T49439 hypothetical prote
108 51 76.1 448 2 S33926 Wilms' tumor prote
109 51 76.1 449 2 A38080 Wilms tumor suscep
110 51 76.1 449 2 A39692 Wilms' tumor prote
111 51 76.1 485 2 A33647 sulfated surface g
112 51 76.1 516 2 T00974 probable SF16 prot
113 51 76.1 530 2 A45690 transactivator EBN
114 51 76.1 571 2 C75530 conserved hypothet
115 51 76.1 599 2 T10798 pherophorin-S - Vo
116 51 76.1 610 2 T06690 galactonolactone d
117 51 76.1 715 1 TNBE77 77K alpha trans-in
118 51 76.1 760 2 T06291 extensin homolog T
119 51 76.1 814 2 T49207 receptor kinase-li
120 51 76.1 1006 2 G86292 hypothetical prote
121 51 76.1 1248 2 T05059 hypothetical prote
122 51 76.1 1392 2 T51947 probable transcrip
123 51 76.1 1737 2 T00209 MEGF8 protein - hu
124 50.5 75.4 437 2 T14192 extensin homolog T
125 50 74.6 141 2 A34043 hypothetical proli
126 50 74.6 142 2 B41132 collagen-related p
127 50 74.6 148 2 S39206 proline-rich prote
128 50 74.6 161 2 H85120 probable proline-r
129 50 74.6 161 2 F86364 hypothetical prote
130 50 74.6 165 2 C71717 hypothetical prote
131 50 74.6 182 2 A44157 spermatophorin Sp2
132 50 74.6 195 2 T07735 nodulin T24A6.3 [i
133 50 74.6 209 2 C89005 proline-rich prote
134 50 74.6 219 2 T10563 nodulin-30 (Npv30)
135 50 74.6 220 2 S52005 TonB protein XF000
136 50 74.6 221 2 H82857 probable AP2 domai
137 50 74.6 236 2 T02577 hypothetical prote
138 50 74.6 246 2 T46446 homeotic protein H
139 50 74.6 250 1 A31757 homeotic protein H
140 50 74.6 251 1 B60492 homeotic protein H
141 50 74.6 255 2 C36222 phosphoprotein pho
142 50 74.6 262 2 T06600 acetyl-CoA carboxy
143 50 74.6 294 2 A55477 survival motor neu
144 50 74.6 301 2 T50001 hypothetical prote
145 50 74.6 302 2 C86480 33.2K hypothetical
146 50 74.6 309 1 S09257 homeotic protein H
147 50 74.6 314 2 D71410 hypothetical prote
148 50 74.6 328 2 C42756 hypothetical prote
149 50 74.6 352 2 C96643 hypothetical prote
150 50 74.6 357 2 T47367 CCAAT/enhancer-bin
151 50 74.6 358 2 A54265 CCAAT/enhancer-bin
152 50 74.6 363 2 T16755 hypothetical prote
153 50 74.6 365 2 T24955 hypothetical prote
154 50 74.6 383 2 AE2295 hypothetical prote
155 50 74.6 388 2 T26922 hypothetical prote
156 50 74.6 389 2 T03411 protein RFL - rice
157 50 74.6 395 2 I49575 CCAAT/enhancer bin
158 50 74.6 401 2 A48423 engrailed homeodom
159 50 74.6 415 1 A34170 acrosin (EC 3.4.21
160 50 74.6 416 2 T34279 hypothetical prote
161 50 74.6 431 2 S47538 acrosin (EC 3.4.21
162 50 74.6 432 2 AB2222 twitching motility
163 50 74.6 436 2 B55452 cartilage-derived
164 50 74.6 440 2 I49681 glyceraldehyde-3-p
165 50 74.6 440 2 S71795 transcription fact
166 50 74.6 445 2 S19990 alpha-amylase (EC
167 50 74.6 446 2 A34418 H-2 region II bind
168 50 74.6 448 2 D41727 retinoid X recepto
169 50 74.6 449 2 D87682 OmpA family protei
170 50 74.6 451 2 A41651 retinoic acid rece
171 50 74.6 457 2 E88456 protein W03A5.3 [i
172 50 74.6 477 2 JC4386 adenyl cyclase-a
173 50 74.6 483 2 D87752 protein Cl8E3.6 [i
174 50 74.6 483 2 T15180 hypothetical prote
175 50 74.6 497 2 T27012 hypothetical prote

176 50 74.6 502 2 A55197 Wiskott-Aldrich sy
177 50 74.6 505 2 S72273 actin-depolymerizi
178 50 74.6 512 2 S28267 pherophorin I prec
179 50 74.6 514 1 A36222 phosphoprotein pho
180 50 74.6 515 2 JT0976 phosphoprotein pho
181 50 74.6 520 2 I84718 RXR-beta1 isoform
182 50 74.6 520 2 G86414 probable protein k
183 50 74.6 524 1 B36222 phosphoprotein pho
184 50 74.6 525 2 A33794 phosphoprotein pho
185 50 74.6 533 2 S37781 retinoid X recepto
186 50 74.6 555 2 T30349 structural protein
187 50 74.6 558 2 G96522 F11A17.16 [importe
188 50 74.6 561 2 S17821 protoporphyirin IX
189 50 74.6 561 2 G96752 unknown protein F2
190 50 74.6 574 2 T05964 probable low-affin
191 50 74.6 581 2 E86408 F3H9.11 protein -
192 50 74.6 589 2 T05074 hypothetical prote
193 50 74.6 600 2 T14463 galactonolactone d
194 50 74.6 602 2 T45760 hypothetical prote
195 50 74.6 622 2 T15467 hypothetical prote
196 50 74.6 637 2 T04552 hypothetical prote
197 50 74.6 666 2 T31461 probable magnesium
198 50 74.6 678 2 AF1841 Mg chelataase chain
199 50 74.6 707 2 S60588 drebrin A - rat
200 50 74.6 742 2 F84643 hypothetical prote
201 50 74.6 754 2 A85043 probable LRR recep
202 50 74.6 760 2 B96724 hypothetical prote
203 50 74.6 783 1 A38637 Ras interactor RIN
204 50 74.6 825 2 S54465 YTA12 protein prec
205 50 74.6 847 2 F96531 hypothetical prote
206 50 74.6 880 2 T48477 hypothetical prote
207 50 74.6 881 2 T33810 hypothetical prote
208 50 74.6 910 2 A34721 hypothetical prote
209 50 74.6 911 2 B34721 androgen receptor
210 50 74.6 919 2 A39248 androgen receptor
211 50 74.6 980 2 S54986 regulatory protein
212 50 74.6 1012 2 T13603 probable N-methyl-
213 50 74.6 1013 2 T33470 hypothetical prote
214 50 74.6 1095 2 T20528 hypothetical prote
215 50 74.6 1100 2 JC8033 leukocyte formin p
216 50 74.6 1110 2 T19673 hypothetical prote
217 50 74.6 1173 2 T31421 C-terminal domain-
218 50 74.6 1176 2 T49482 hypothetical prote
219 50 74.6 1213 2 A41724 limb deformity (ld
220 50 74.6 1264 2 A36858 G2R protein - vari
221 50 74.6 1443 2 S05979 steroid hormone re
222 50 74.6 1541 2 T02831 AAA protein L4171.
223 50 74.6 1584 2 T00026 brain-specific ang
224 50 74.6 1607 2 T03022 MAP kinase kinase
225 50 74.6 1621 2 T15264 hypothetical prote
226 50 74.6 1684 2 T02632 hypothetical prote
227 50 74.6 1896 2 B72175 D15R protein - var
228 50 74.6 1897 2 T28621 hypothetical prote
229 50 74.6 1952 2 T48814 hypothetical prote
230 50 74.6 2649 2 T51023 hypothetical prote
231 50 74.6 3119 2 I49729 HD protein - mouse
232 50 74.6 3144 2 A46068 Huntington disease
233 49 73.1 57 2 S10782 salivary protein P
234 49 73.1 67 2 S14971 extensin class I (
235 49 73.1 75 2 S14973 extensin class I (
236 49 73.1 79 1 PJHUSB proline-rich pepti
237 49 73.1 90 2 S14972 extensin class I (
238 49 73.1 117 2 T46322 hypothetical prote
239 49 73.1 132 2 S14970 extensin class I (
240 49 73.1 163 2 A29356 hydroxyproline-ric
241 49 73.1 169 2 T07623 extensin homolog H
242 49 73.1 169 2 T34520 hypothetical prote
243 49 73.1 181 2 S14974 extensin class I (
244 49 73.1 230 2 T29591 hypothetical prote
245 49 73.1 242 2 S54156 extensin-like prot
246 49 73.1 270 2 S74993 hypothetical prote
247 49 73.1 279 2 T02495 hypothetical prote
248 49 73.1 280 2 T11671 extensin-like prot

249 49 73.1 287 2 T34397 hypothetical prote
250 49 73.1 358 2 E87309 hypothetical prote
251 49 73.1 368 2 C29356 hydroxyproline-ric
252 49 73.1 368 2 T51200 hypothetical prote
253 49 73.1 379 2 S31719 proline-rich prote
254 49 73.1 388 2 S25298 extensin (clone To
255 49 73.1 440 2 H86312 F2H15.2 prote in -
256 49 73.1 440 2 JC7807 Wiskott-Aldrich sy
257 49 73.1 445 2 S41741 N-acetylmuramoyl-L
258 49 73.1 445 2 A86113 N-acetylmuramoyl-L
259 49 73.1 445 2 A91272 N-acetylmuramoyl-L
260 49 73.1 459 2 T31608 hypothetical prote
261 49 73.1 483 2 F71684 virB10 prote in (vi
262 49 73.1 485 2 T27147 hypothetical prote
263 49 73.1 485 2 PC4427 Nck, Ash and phosp
264 49 73.1 489 2 T11622 extensin class 1 p
265 49 73.1 496 2 T48812 hypothetical prote
266 49 73.1 530 2 T50498 myc-like prote in -
267 49 73.1 551 2 S57447 HPBRII-7 prote in -
268 49 73.1 577 2 T09024 proline-rich prote
269 49 73.1 583 1 A41129 radixin - mouse
270 49 73.1 583 1 A46127 radixin - human
271 49 73.1 583 1 S39805 radixin - pig
272 49 73.1 603 2 F96797 probable GT-like t
273 49 73.1 613 2 T00077 gag-like prote in -
274 49 73.1 639 2 G02919 transcription fact
275 49 73.1 650 2 B87791 protein B0207.1 [i
276 49 73.1 697 2 C75525 molybdopterin oxid
277 49 73.1 710 2 D96728 hypothetical prote
278 49 73.1 734 2 T04876 hypothetical prote
279 49 73.1 745 2 T51370 hypothetical prote
280 49 73.1 786 2 T01456 extensin homolog F
281 49 73.1 917 2 T21870 hypothetical prote
282 49 73.1 1174 2 T38140 protein-tyrosine-p
283 49 73.1 1268 2 T31420 C-terminal domain-
284 49 73.1 1410 2 T47137 hypothetical prote
285 49 73.1 1560 2 T00080 hypothetical prote
286 49 73.1 1575 2 S68448 synaptojanin, 170K
287 49 73.1 1802 2 H88444 protein C26E6.12 [i
288 49 73.1 3005 2 T08841 polyprote in - dour
289 48.5 72.4 225 2 T09964 extensin CYC15 pre
290 48.5 72.4 240 2 D70894 probable pra prote
291 48.5 72.4 249 2 A41497 36K antigen pra -
292 48.5 72.4 443 2 G86388 unknown prote in [i
293 48.5 72.4 592 2 D70863 hypothetical prote
294 48 71.6 100 2 T17126 hypothetical prote
295 48 71.6 135 2 E96750 hypothetical prote
296 48 71.6 185 2 B26669 nodulin-20 precurs
297 48 71.6 207 2 I53154 scleraxis - mouse
298 48 71.6 208 2 A43696 Spec3 prote in - se
299 48 71.6 217 2 T09965 extensin CYC17 pre
300 48 71.6 323 2 T49723 hypothetical prote
301 48 71.6 338 1 TVMSFB transforming prote
302 48 71.6 338 2 I53043 transforming prote
303 48 71.6 350 2 T14191 extensin homolog T
304 48 71.6 356 1 WJHU2H homeotic prote in H
305 48 71.6 371 1 FOMVCS gag polyprote in -
306 48 71.6 384 2 H88924 prote in F33E11.2 [i
307 48 71.6 392 2 B48423 homeotic prote in e
308 48 71.6 424 2 A54964 spliceosome-associ
309 48 71.6 435 2 A42672 choline kinase (EC
310 48 71.6 453 2 JX0342 choline kinase (EC
311 48 71.6 474 2 I49572 adenylyl cyclase-a
312 48 71.6 476 2 T46067 hypothetical prote
313 48 71.6 536 1 FOMVMD gag polyprote in -
314 48 71.6 547 2 C96828 unknown prote in F1
315 48 71.6 550 2 G70597 probable prote inas
316 48 71.6 557 2 A55933 paxillin - human
317 48 71.6 559 2 B55933 paxillin - chicken
318 48 71.6 569 2 A46462 T cell activation
319 48 71.6 575 2 C88346 prote in F42G4.3a [i
320 48 71.6 594 2 B86456 prote in trihelix D
321 48 71.6 603 2 T22111 hypothetical prote

322 48 71.6 653 2 T47581
323 48 71.6 653 2 T51244
324 48 71.6 678 2 B70913
325 48 71.6 736 2 I51691
326 48 71.6 760 2 F86387
327 48 71.6 790 1 FOMVHZ
328 48 71.6 1204 2 T19918
329 48 71.6 1483 2 E86143
330 48 71.6 1607 2 T21982
331 48 71.6 1615 2 JE0372
332 48 71.6 1794 2 T38459
333 48 71.6 1841 2 T38091
334 48 71.6 1872 2 T24683
335 48 71.6 3848 2 T17414
336 47.5 70.9 1409 2 T37188
337 47 70.1 30 2 S59482
338 47 70.1 164 2 T26608
339 47 70.1 199 2 T07622
340 47 70.1 212 2 F84862
341 47 70.1 215 2 S34163
342 47 70.1 217 2 S09754
343 47 70.1 217 2 T15873
344 47 70.1 219 2 I51382
345 47 70.1 226 2 S41032
346 47 70.1 229 1 W4WL8
347 47 70.1 240 2 T25814
348 47 70.1 245 1 W4WL5
349 47 70.1 245 1 W4WLB5
350 47 70.1 258 2 S55036
351 47 70.1 264 2 PQ0478
352 47 70.1 281 2 S65692
353 47 70.1 301 2 JQ1663
354 47 70.1 303 1 TVHUJD
355 47 70.1 304 1 W4WL47
356 47 70.1 320 2 T18319
357 47 70.1 336 2 S29507
358 47 70.1 345 2 T01348
359 47 70.1 347 2 A43815
360 47 70.1 348 2 D88088
361 47 70.1 348 2 S52720
362 47 70.1 411 2 S34164
363 47 70.1 417 2 S47539
364 47 70.1 432 2 T06782
365 47 70.1 463 2 T45565
366 47 70.1 467 2 T34874
367 47 70.1 474 2 A46584
368 47 70.1 474 2 T10271
369 47 70.1 508 2 T45867
370 47 70.1 517 2 E89530
371 47 70.1 546 2 T02029
372 47 70.1 580 2 T10863
373 47 70.1 585 2 I58403
374 47 70.1 593 2 S15215
375 47 70.1 658 2 T08153
376 47 70.1 707 2 A46302
377 47 70.1 810 2 C70791
378 47 70.1 894 2 F84870
379 47 70.1 927 2 A48085
380 47 70.1 969 2 T15446
381 47 70.1 997 2 T28872
382 47 70.1 1003 2 T34066
383 47 70.1 1010 2 A33509
384 47 70.1 1032 2 D83637
385 47 70.1 1048 2 T31425
386 47 70.1 1066 2 T10108
387 47 70.1 1100 2 T30967
388 47 70.1 1134 1 A35955
389 47 70.1 1135 1 A29997
390 47 70.1 1252 2 T14272
391 47 70.1 1311 2 T33757
392 47 70.1 1400 2 T52359
393 47 70.1 1420 2 T37781
394 47 70.1 1501 1 B29813

SCARECROW1 - Arabi
SCARECROW prote in
probable penicilli
dishevelled homolo
probable Pto kinas
gag-kit polyprotei
hypothetical prote
F6F3.12 prote in -
hypothetical prote
low density lipopr
hypothetical diver
cell division cont
hypothetical prote
TipC prote in - sli
presynaptic activi
hydroxyproline-ric
hypothetical prote
extensin homolog -
probable extensin
homeotic prote in H
hypothetical prote
hypothetical prote
achaete-scute homo
hypothetical prote
E4 prote in - human
hypothetical prote
E4 prote in - human
E4 prote in - human
tyrosine-rich hydr
pistil extensin-li
E4 prote in - human
hybrid proline-ric
transforming prote
E4 prote in - human
hypothetical prote
metavinculin - pig
hypothetical prote
transforming prote
prote in B0454.1 [i
homeobox prote in g
homeotic prote in H
homeotic prote in H
extensin - soybean
hypothetical prote
hypothetical prote
adenylyl cyclase-a
capsid-associated
hypothetical prote
prote in H28G03.2 [i
DNA-binding protei
extensin precursor
H4 prote in - human
virulence-associat
cysteine prote inas
PTB-associated spl
probable pona' pro
hypothetical prote
transcription fact
hypothetical prote
hypothetical prote
hypothetical prote
vinculin - Caenorh
serine/threonine p
C-terminal domain-
vinculin - mouse
transcription acti
meta-vinculin - hu
meta-vinculin - ch
cortactin-binding
hypothetical prote
hypothetical prote
probable cytoskele
174K ninaC prote in

395 47 70.1 1584 2 T18276 protein-tyrosine k
396 47 70.1 3124 2 A40020 collagen alpha 1(X
397 46.5 69.4 306 2 A24354 extensin precursor
398 46.5 69.4 366 2 T26449 hypothetical prote
399 46.5 69.4 419 2 G70602 hypothetical prote
400 46.5 69.4 457 2 I55976 dihydrolipoamide S
401 46.5 69.4 512 1 FOMVGS gag polyprotein -
402 46.5 69.4 520 1 FOLJGL gag polyprotein -
403 46.5 69.4 555 2 S21766 dihydrolipoamide S
404 46.5 69.4 1464 2 T13716 bazooka gene prote
405 46.5 69.4 1466 2 T17138 CL1AA protein - ra
406 46.5 69.4 1467 2 T18411 latrophilin-1, bra
407 46.5 69.4 1471 2 T17149 CL1BA protein - ra
408 46.5 69.4 1472 2 T18413 latrophilin-1, bra
409 46.5 69.4 1510 2 T17145 CL1AB protein - ra
410 46.5 69.4 1515 2 T17156 CL1BB protein - ra
411 46 68.7 80 2 PQ0477 pistil extensin-li
412 46 68.7 90 2 T18068 hypothetical prote
413 46 68.7 91 2 S37486 gene msg3 protei
414 46 68.7 140 2 T19083 hypothetical prote
415 46 68.7 147 2 S37485 gene msg1 protei
416 46 68.7 153 2 T31654 hypothetical prote
417 46 68.7 154 2 PQ0476 pistil extensin-li
418 46 68.7 156 2 C84556 probable bZIP tran
419 46 68.7 172 2 T27505 hypothetical prote
420 46 68.7 173 2 T51469 glycine/proline-ri
421 46 68.7 194 2 A38203 proline-rich prote
422 46 68.7 204 2 T07679 protein import rec
423 46 68.7 222 2 H96711 hypothetical prote
424 46 68.7 244 2 S44822 F4482.3 protei
425 46 68.7 248 2 F83558 hypothetical prote
426 46 68.7 250 2 T08908 hypothetical prote
427 46 68.7 253 2 T17312 hypothetical prote
428 46 68.7 264 2 B89005 protein T24A6.18 [
429 46 68.7 275 2 T02559 probable spliceoso
430 46 68.7 283 2 S13383 hydroxyproline-ric
431 46 68.7 322 2 T22403 hypothetical prote
432 46 68.7 325 2 D70728 hypothetical prote
433 46 68.7 340 2 S53763 homeotic protein H
434 46 68.7 345 2 S12788 transcription fact
435 46 68.7 346 2 T27896 hypothetical prote
436 46 68.7 369 2 S20500 hydroxyproline-ric
437 46 68.7 370 2 B88455 protein T15B12.2 [
438 46 68.7 373 2 A70856 probable lpp2 prot
439 46 68.7 377 2 JC4368 protein farnesyltr
440 46 68.7 377 2 A41625 protein farnesyltr
441 46 68.7 379 2 T19069 hypothetical prote
442 46 68.7 380 2 S51797 vasodilator-stimul
443 46 68.7 384 2 S51796 vasodilator-stimul
444 46 68.7 387 2 T21370 hypothetical prote
445 46 68.7 388 2 JC5437 spliceosome-associ
446 46 68.7 393 2 T33103 lin-1 protein - Ca
447 46 68.7 407 2 H87341 OmpA family protei
448 46 68.7 420 2 T46910 hypothetical prote
449 46 68.7 427 2 T48159 hypothetical prote
450 46 68.7 432 2 B96515 hypothetical prote
451 46 68.7 447 2 S37048 cysteine proteinas
452 46 68.7 447 2 A84718 hypothetical prote
453 46 68.7 450 2 S07051 cysteine proteinas
454 46 68.7 450 2 S12099 cysteine proteinas
455 46 68.7 461 1 A46077 steroid hormone re
456 46 68.7 462 1 A56120 steroid hormone re
457 46 68.7 462 2 A40716 steroid hormone re
458 46 68.7 464 2 JC7143 endoglucanase I -
459 46 68.7 465 2 A42128 embryonal long ter
460 46 68.7 475 2 A48120 adenylyl cyclase-a
461 46 68.7 480 2 C86357 hypothetical prote
462 46 68.7 487 2 S65310 probable membrane
463 46 68.7 491 2 S67694 probable membrane
464 46 68.7 496 2 S26402 homeotic protein H
465 46 68.7 498 2 C72167 A30L protein - var
466 46 68.7 498 2 C36851 A29L protein - var
467 46 68.7 498 2 T28571 hypothetical prote

468 468 68.7 502 2 T08776 hypothetical prote
469 46 68.7 514 2 E83644 chromosomal replic
470 46 68.7 538 1 FOMVLM gag polypotein -
471 46 68.7 541 2 T19304 hypothetical prote
472 46 68.7 542 2 T29707 hypothetical prote
473 46 68.7 544 2 T17547 proline-rich prote
474 46 68.7 551 2 A60047 adenylyl cyclase-a
475 46 68.7 554 2 T49833 related to VEA pro
476 46 68.7 556 2 D70940 probable ppg prote
477 46 68.7 558 2 A33616 heterogeneous ribo
478 46 68.7 559 2 B44265 ENL (translocation
479 46 68.7 564 2 T45866 hypothetical prote
480 46 68.7 581 2 S17150 potassium channel
481 46 68.7 581 2 T34478 hypothetical prote
482 46 68.7 612 2 T02414 probable protein k
483 46 68.7 613 2 A39402 potassium channel
484 46 68.7 624 2 S22703 voltage-gated pota
485 46 68.7 628 2 D86466 69.4K hypothetical
486 46 68.7 633 2 T34479 hypothetical prote
487 46 68.7 643 2 T27429 hypothetical prote
488 46 68.7 650 2 S72367 ATP-dependent RNA
489 46 68.7 655 2 AD2422 DNA polymerase III
490 46 68.7 658 2 T04831 probable serine/th
491 46 68.7 667 2 T17221 hypothetical prote
492 46 68.7 684 2 A56154 Abl substrate ena
493 46 68.7 695 2 T24950 hypothetical prote
494 46 68.7 739 2 A88445 protein C26E6.10 [
495 46 68.7 754 2 T06249 protoporphylin IX
496 46 68.7 758 2 T02925 protoporphylin IX
497 46 68.7 778 2 B86218 protein T27G7.20 [
498 46 68.7 798 2 T19864 hypothetical prote
499 46 68.7 798 2 T21369 hypothetical prote
500 46 68.7 814 2 T05537 probable serine/th
501 46 68.7 817 2 S51342 verprolin - yeast
502 46 68.7 856 2 T13159 E1B-55kDa-associat
503 46 68.7 895 2 C86371 99.7K hypothetical
504 46 68.7 899 2 T16204 hypothetical prote
505 46 68.7 921 2 A33718 retinoblastoma pro
506 46 68.7 948 2 A57640 retinoblastoma bin
507 46 68.7 981 1 FOMVGM gag-abl polypotei
508 46 68.7 984 2 A29513 mineralocorticoid
509 46 68.7 1018 2 T43168 hypothetical prote
510 46 68.7 1048 2 T23764 hypothetical prote
511 46 68.7 1133 2 A54164 sterol regulatory
512 46 68.7 1139 2 A49370 E1A-associated cyc
513 46 68.7 1148 2 T09073 splicing factor Si
514 46 68.7 1217 2 T39427 probable myosin I
515 46 68.7 1219 2 S54570 probable membrane
516 46 68.7 1259 2 T16038 hypothetical prote
517 46 68.7 1296 2 T13936 collar protein iso
518 46 68.7 1299 1 S06119 membrane protein p
519 46 68.7 1335 2 T18289 racGAP protein - s
520 46 68.7 1366 2 B84924 hypothetical prote
521 46 68.7 1445 2 A59437 KIAA1204 protein [
522 46 68.7 1461 2 T41643 probable involveme
523 46 68.7 1560 2 T42727 proliferation pote
524 46 68.7 1663 2 T42092 s-afadin - rat
525 46 68.7 1829 2 T41751 l-afadin - rat
526 46 68.7 1882 1 GNVVTR genome polyprotein
527 46 68.7 2363 2 T38841 probable pre-mRNA
528 46 68.7 2413 2 S34670 splicing factor PR
529 45.5 67.9 74 2 B40513 hypothetical prote
530 45.5 67.9 135 2 A88951 protein C38C3.6 [i
531 45.5 67.9 198 2 A45067 laminin B1 chain v
532 45.5 67.9 858 2 JG0183 myosin Myok - Dict
533 45.5 67.9 1151 2 T18535 high molecular mas
534 45.5 67.9 1402 2 T17456 cell surface prote
535 45 67.2 75 2 T26916 hypothetical prote
536 45 67.2 101 2 T46506 hypothetical prote
537 45 67.2 115 2 T36886 hypothetical prote
538 45 67.2 132 2 H96708 hypothetical prote
539 45 67.2 154 2 T41831 AcMNPV orf91 - Bom
540 45 67.2 190 2 T35570 hypothetical prote

541	45	67.2	194	2	T12504	hypothetical prote	614	44.5	66.4	164	2	T15525	hypothetical prote
542	45	67.2	212	2	S74288	hypothetical prote	615	44.5	66.4	172	2	T23261	hypothetical prote
543	45	67.2	214	2	T10737	extensin-like cell	616	44.5	66.4	316	2	T31880	hypothetical prote
544	45	67.2	214	2	T09854	proline-rich cell	617	44.5	66.4	357	2	PC4293	nuclear factor 1 f
545	45	67.2	220	2	S42879	nodulin-30 - kidne	618	44.5	66.4	461	2	S45568	nuclear factor I-A
546	45	67.2	269	2	T30468	hypothetical prote	619	44.5	66.4	491	2	T07598	proline-rich prote
547	45	67.2	274	2	T04619	hypothetical prote	620	44.5	66.4	498	2	S45567	nuclear factor I-A
548	45	67.2	281	2	T29150	hypothetical prote	621	44.5	66.4	509	2	S45565	nuclear factor I-A
549	45	67.2	301	2	T18788	hypothetical prote	622	44.5	66.4	509	2	S45565	nuclear factor I-A
550	45	67.2	313	1	FOVDA	gag polyprotein -	623	44.5	66.4	513	2	T14194	extensin homolog T
551	45	67.2	320	2	AE2842	conserved hypothet	624	44.5	66.4	522	2	S09996	nuclear factor I-A
552	45	67.2	332	2	G86182	hypothetical prote	625	44.5	66.4	531	2	G85071	hypothetical prote
553	45	67.2	336	2	E47301	Vir10 homolog - Bo	626	44.5	66.4	532	2	B36596	nuclear factor I -
554	45	67.2	346	2	S76923	hypothetical prote	627	44.5	66.4	688	2	T09941	transcription fact
555	45	67.2	358	2	T01296	leucine-rich repea	628	44.5	66.4	719	2	B86490	F28L22.6 protein -
556	45	67.2	359	2	F97619	hypothetical prote	629	44.5	66.4	950	2	T22592	hypothetical prote
557	45	67.2	367	2	I39172	cyclin-dependent k	630	44.5	66.4	1274	2	T37193	enamelin matrix pr
558	45	67.2	373	2	A47234	homeobox protein H	631	44.5	66.4	2957	2	T33152	hypothetical prote
559	45	67.2	386	2	E72645	hypothetical prote	632	44	65.7	17	2	S59481	hydroxyproline-ric
560	45	67.2	388	2	E72103	hypothetical prote	633	44	65.7	88	4	S00849	hypothetical prote
561	45	67.2	388	2	C81568	hypothetical prote	634	44	65.7	98	2	T02437	hydroxyproline-ric
562	45	67.2	388	2	H86520	hypothetical prote	635	44	65.7	134	2	JC5572	hypothetical trans
563	45	67.2	389	2	B82819	outer membrane pro	636	44	65.7	137	1	SSKV	proline-rich prote
564	45	67.2	405	2	T42663	hypothetical prote	637	44	65.7	139	2	T05847	cupredoxin [valida
565	45	67.2	416	2	S27198	homeotic protein H	638	44	65.7	141	2	T06646	hypothetical prote
566	45	67.2	416	2	JU0465	extensin precursor	639	44	65.7	155	2	T16855	blue copper-bindin
567	45	67.2	427	2	A53798	58K membrane-assoc	640	44	65.7	156	1	GNVQL2	hypothetical prote
568	45	67.2	439	2	S51939	chitinase (EC 3.2.	641	44	65.7	156	1	GNVQLL	genome-linked prot
569	45	67.2	440	2	S51614	Algal-CAM - Volvox	642	44	65.7	156	1	GNVQWA	genome-linked prot
570	45	67.2	472	2	C70853	hypothetical prote	643	44	65.7	164	2	H87551	conserved hypothet
571	45	67.2	473	2	T05306	homeobox protein A	644	44	65.7	172	2	A39458	carotene biosynthe
572	45	67.2	474	2	A36240	pyrimidine synthes	645	44	65.7	178	2	S26044	sex-determining pr
573	45	67.2	474	2	S57239	gene forked protei	646	44	65.7	181	2	S53178	core antigen - hep
574	45	67.2	491	2	S76943	hypothetical prote	647	44	65.7	183	2	S53155	core antigen - hep
575	45	67.2	495	1	S31223	transcription fact	648	44	65.7	183	2	S53184	core antigen - hep
576	45	67.2	515	1	T38946	phosphoprotein pho	649	44	65.7	184	2	S26046	sex-determining pr
577	45	67.2	535	2	T45831	SF16-like protein	650	44	65.7	185	2	S53288	core antigen - hep
578	45	67.2	544	2	S44814	F44B9.3 protein -	651	44	65.7	191	2	E87601	OmpA family protei
579	45	67.2	578	2	AG0119	probable OmpA-fami	652	44	65.7	197	2	A29648	female-specific tr
580	45	67.2	578	2	AC0179	probable exported	653	44	65.7	198	2	T35752	hypothetical prote
581	45	67.2	604	2	S39885	forked protein - f	654	44	65.7	211	2	B89716	protein F45B8.3 [i
582	45	67.2	627	2	T26064	hypothetical prote	655	44	65.7	212	2	S53157	e antigen precursor
583	45	67.2	641	2	S57236	forked protein 2.5	656	44	65.7	218	2	F84748	probable AP2 domai
584	45	67.2	650	2	T22002	hypothetical prote	657	44	65.7	226	2	T35236	hypothetical prote
585	45	67.2	655	1	A55726	RNA-binding protei	658	44	65.7	228	2	S53504	extensin-like prot
586	45	67.2	656	1	A49358	RNA-binding protei	659	44	65.7	236	2	T45835	hypothetical prote
587	45	67.2	661	2	T15073	hypothetical prote	660	44	65.7	241	2	T22216	hypothetical prote
588	45	67.2	684	2	T01267	leucine-rich repea	661	44	65.7	279	2	A53062	Fas ligand - mouse
589	45	67.2	696	2	T31538	hypothetical prote	662	44	65.7	281	2	D70845	hypothetical prote
590	45	67.2	699	2	A96529	hypothetical prote	663	44	65.7	283	2	G01926	insulin promoter f
591	45	67.2	708	2	D96711	hypothetical prote	664	44	65.7	291	2	S33209	extensin-like prot
592	45	67.2	711	2	S68443	double-stranded RN	665	44	65.7	299	2	G70784	probable mmpS3 pro
593	45	67.2	721	2	E70766	hypothetical prote	666	44	65.7	303	2	JQ1386	hypothetical 33K p
594	45	67.2	725	2	T01268	leucine-rich repea	667	44	65.7	304	2	T48281	hypothetical prote
595	45	67.2	745	2	S13586	triacylglycerol li	668	44	65.7	306	2	T09067	extensin-like prot
596	45	67.2	751	2	T34490	hypothetical prote	669	44	65.7	309	2	B87576	hypothetical prote
597	45	67.2	753	2	A27041	tyrosine kinase-re	670	44	65.7	311	2	T02783	probable homeotic
598	45	67.2	772	2	T13078	KIAA0992 protein -	671	44	65.7	324	2	F84913	probable FCA-relat
599	45	67.2	906	2	A71438	probable resistanc	672	44	65.7	324	2	T24102	hypothetical prote
600	45	67.2	954	1	S68178	mixed-lineage prot	673	44	65.7	331	2	B47236	zinc-finger protei
601	45	67.2	986	2	T33135	hypothetical prote	674	44	65.7	348	2	A49594	enhancer factor pr
602	45	67.2	1076	2	S50536	hypothetical prote	675	44	65.7	348	2	T47494	hypothetical prote
603	45	67.2	1099	2	A56155	tumor suppressor p	676	44	65.7	352	2	T18794	hypothetical prote
604	45	67.2	1110	2	A43253	Large tra-1 protei	677	44	65.7	370	2	JC2204	MAPK-activated pro
605	45	67.2	1196	2	T23832	protein-tyrosine k	678	44	65.7	375	2	T51333	transcription fact
606	45	67.2	1403	2	S24548	homeotic protein p	679	44	65.7	381	2	S52985	cell wall protein
607	45	67.2	1436	2	S57238	forked protein 5.4	680	44	65.7	392	1	F0LJGB	gag polyprotein -
608	45	67.2	1449	2	S57237	forked protein 5.6	681	44	65.7	393	2	S29356	gag protein - bovi
609	45	67.2	2205	1	MNWVRN	nonstructural poly	682	44	65.7	396	2	S39793	MAPK-activated pro
610	45	67.2	2212	2	T28157	erythrocyte membra	683	44	65.7	396	2	A36339	FLO protein - gard
611	45	67.2	2871	2	A55624	fibillin-1 precur	684	44	65.7	399	2	I49754	homeobox protein -
612	45	67.2	3511	2	A59295	unconventional myo	685	44	65.7	401	2	T24381	hypothetical prote
613	44.5	66.4	139	2	S61885	extensin precursor	686	44	65.7	408	2	G96707	hypothetical prote

687 44 65.7 411 2 A70509 hypothetical prote
688 44 65.7 417 2 G64417 hypothetical prote
689 44 65.7 418 2 F82638 hypothetical prote
690 44 65.7 419 2 T29266 hypothetical prote
691 44 65.7 425 1 F0MVG C gag polyprotei
692 44 65.7 426 2 JQ1696 pistil extensi
693 44 65.7 429 2 T06296 extensin-like
694 44 65.7 433 1 F0LJH2 gag polyprotei
695 44 65.7 434 2 C96515 hypothetical prote
696 44 65.7 447 2 S44809 F44B9.8 protei
697 44 65.7 448 2 T06076 proline-rich prote
698 44 65.7 449 2 H70652 hypothetical prote
699 44 65.7 451 2 T30603 perlecan homolog
700 44 65.7 461 2 T10741 extensin-like
701 44 65.7 473 2 E75097 hypothetical prote
702 44 65.7 477 2 A47236 zinc-finger protei
703 44 65.7 478 2 C29514 muscarinic acetyl
704 44 65.7 478 2 F86388 hypothetical prote
705 44 65.7 479 2 S10127 muscarinic acetyl
706 44 65.7 479 2 S33776 muscarinic acetyl
707 44 65.7 480 2 JH0672 brain factor 1
708 44 65.7 488 2 F86209 protein F22G5.14
709 44 65.7 494 2 A42170 zinc finger protei
710 44 65.7 497 2 JC5076 myc-associated zin
711 44 65.7 526 2 A34896 adenylate cyclase
712 44 65.7 532 2 JC6170 GATA-transcription
713 44 65.7 532 2 T34235 hypothetical prote
714 44 65.7 554 2 T02445 probable U4/U6 sma
715 44 65.7 559 2 G84642 hypothetical prote
716 44 65.7 590 1 WMBECB 64K capsid assembl
717 44 65.7 593 2 T24379 hypothetical prote
718 44 65.7 605 1 Q0BE3R BVRP2 (EC-RF3) pro
719 44 65.7 620 2 S06733 hydroxyproline-ric
720 44 65.7 626 2 H71173 hypothetical prote
721 44 65.7 630 2 T02524 probable RING zinc
722 44 65.7 634 2 T00388 hypothetical prote
723 44 65.7 644 2 S15464 gp70 protein - mur
724 44 65.7 655 2 S40521 FKHR protein - hum
725 44 65.7 664 2 T01368 hypothetical prote
726 44 65.7 670 2 JC5887 signaling mediator
727 44 65.7 679 2 S37842 hypothetical prote
728 44 65.7 694 2 T01005 hypothetical prote
729 44 65.7 707 2 T14195 extensin homolog
730 44 65.7 733 2 T47618 extensin-like prote
731 44 65.7 756 2 JC5886 signaling mediator
732 44 65.7 789 2 T52067 hypothetical prote
733 44 65.7 808 2 C72858 AcOrf-66 protein -
734 44 65.7 824 2 T16028 hypothetical prote
735 44 65.7 868 2 T20239 hypothetical prote
736 44 65.7 882 2 S41034 hypothetical prote
737 44 65.7 899 2 A35895 androgen receptor
738 44 65.7 902 2 B40494 androgen receptor
739 44 65.7 928 1 RBHU retinoblastoma-ass
740 44 65.7 945 2 T00746 hypothetical prote
741 44 65.7 951 2 T47617 extensin-like prote
742 44 65.7 958 2 T13593 hypothetical prote
743 44 65.7 979 2 A35913 regulatory factor
744 44 65.7 1000 2 S44898 ZK1236.3 protein -
745 44 65.7 1001 2 T28897 hypothetical prote
746 44 65.7 1017 2 T15598 hypothetical prote
747 44 65.7 1186 2 T42729 histocompatibility
748 44 65.7 1203 2 T51029 related to pathway
749 44 65.7 1212 2 T42387 histocompatibility
750 44 65.7 1240 2 JC5209 insulin receptor s
751 44 65.7 1241 2 T18311 hypothetical prote
752 44 65.7 1242 2 JS0670 insulin receptor s
753 44 65.7 1262 2 T13353 protein stn-B - fr
754 44 65.7 1306 2 T13592 hypothetical prote
755 44 65.7 1308 2 T15280 hypothetical prote
756 44 65.7 1347 2 T02214 ubiquitous TPR mot
757 44 65.7 1401 2 T02255 probable ubiquitou
758 44 65.7 1411 2 T48529 hypothetical prote
759 44 65.7 1429 2 T13720 gene expanded prot

760 44 65.7 1460 1 EDBEIF immediate-early pr
761 44 65.7 1612 2 T30805 dutt1 protein - mo
762 44 65.7 1651 2 T14160 transmembrane rece
763 44 65.7 1784 2 T10532 gag-pol polyprotei
764 43.5 64.9 42 2 T07030 extensin - tomato
765 43.5 64.9 330 2 T05717 probable extensin
766 43.5 64.9 385 2 S78100 MAPK-activated pro
767 43.5 64.9 578 2 T25647 hypothetical prote
768 43.5 64.9 790 2 F86936 probable integral
769 43.5 64.9 874 2 T15570 hypothetical prote
770 43.5 64.9 913 2 JC5463 alpha-glucosidase
771 43.5 64.9 1250 2 T00454 hypothetical prote
772 43.5 64.9 1634 1 JC5500 phosphoinositide 3
773 43.5 64.9 1776 2 G86280 protein T5E21.13
774 43 64.2 84 2 T09540 proline rich prote
775 43 64.2 158 2 A86452 protein F6N18.7
776 43 64.2 181 2 JC5233 spermatophorin Sp2
777 43 64.2 184 2 S78091 endocuticular prot
778 43 64.2 184 2 B95350 protein [imported
779 43 64.2 192 2 S76867 hypothetical prote
780 43 64.2 192 2 T30477 hypothetical prote
781 43 64.2 197 2 S35252 proline-rich prote
782 43 64.2 212 2 S57330 cathelin-like anti
783 43 64.2 219 2 T27105 hypothetical prote
784 43 64.2 224 2 D72861 AcOrf-91 protein -
785 43 64.2 226 2 T35435 probable integral
786 43 64.2 228 2 S40463 prophenin (PF-2) p
787 43 64.2 239 2 T03078 conserved hypothet
788 43 64.2 239 2 T36995 hypothetical prote
789 43 64.2 254 2 B84901 hypothetical prote
790 43 64.2 257 2 T10586 small nuclear ribo
791 43 64.2 264 2 T06789 hydroxyproline-ric
792 43 64.2 278 2 S20790 extensin - almond
793 43 64.2 278 2 G86912 hypothetical prote
794 43 64.2 278 2 T10019 hypothetical prote
795 43 64.2 298 2 S53761 triose-phosphate i
796 43 64.2 298 2 T18664 hypothetical prote
797 43 64.2 307 2 B72677 hypothetical prote
798 43 64.2 312 2 T46255 hypothetical prote
799 43 64.2 322 2 S25299 hypothetical prote
800 43 64.2 323 2 S20099 extensin precursor
801 43 64.2 329 2 T10064 transforming prote
802 43 64.2 341 1 TVMSJD cytokinin-induced
803 43 64.2 341 2 JC4051 transforming prote
804 43 64.2 346 2 S19129 jun-D protein - ra
805 43 64.2 355 2 S41285 proline-rich prote
806 43 64.2 361 2 S19552 coat protein - swe
807 43 64.2 372 2 T10472 potassium channel
808 43 64.2 379 2 A47659 G-box binding prot
809 43 64.2 381 2 T27806 farnesyl-protein t
810 43 64.2 383 2 T06753 hypothetical prote
811 43 64.2 392 1 F0LJGA zinc finger protei
812 43 64.2 393 2 T43401 gag polyprotein -
813 43 64.2 411 2 B41398 transcription init
814 43 64.2 414 2 T50010 inhibit beta-B cha
815 43 64.2 416 2 B75455 hypothetical prote
816 43 64.2 418 2 T19800 N-acetyl-gamma-glu
817 43 64.2 427 2 I49603 hypothetical prote
818 43 64.2 430 2 JC2301 transcription regu
819 43 64.2 444 2 E83802 hypothetical prote
820 43 64.2 445 2 T30604 hypothetical prote
821 43 64.2 448 1 A56018 transcription fact
822 43 64.2 449 1 S30205 probable membrane
823 43 64.2 450 2 A86919 transcription fact
824 43 64.2 451 1 A40168 transcription fact
825 43 64.2 467 2 S41318 hypothetical prote
826 43 64.2 468 2 S44815 F44B9.4 protein -
827 43 64.2 469 2 I37451 HBF-G2 (HFK-2) pro
828 43 64.2 476 2 F87324 hypothetical prote
829 43 64.2 476 2 A54743 transcription fact
830 43 64.2 477 2 I38409 adenyl cyclase-a
831 43 64.2 483 2 T25992 hypothetical prote
832 43 64.2 484 2 T07675 cyclin a2-type, mi

833	43	64.2	486	2	A41537	DNA-binding protei	906	42.5	63.4	176	2	A86441	hypothetical prote
834	43	64.2	487	1	S52261	NADH2 dehydrogenas	907	42.5	63.4	279	2	T05421	hypothetical prote
835	43	64.2	488	2	JC7995	transient receptor	908	42.5	63.4	403	2	S52796	prpL2 protein - hu
836	43	64.2	493	2	T01206	cysteine proteinas	909	42.5	63.4	440	1	S60755	rhodopsin - Allote
837	43	64.2	493	2	T29030	hypothetical prote	910	42.5	63.4	461	2	S34472	MFH-1 protein - mo
838	43	64.2	514	2	T15338	hypothetical prote	911	42.5	63.4	462	2	E70955	hypothetical prote
839	43	64.2	551	2	T16557	hypothetical prote	912	42.5	63.4	517	1	ERADA7	early E2A DNA-bind
840	43	64.2	557	2	S62522	nuclear protein SP	913	42.5	63.4	554	2	F86244	hypothetical prote
841	43	64.2	560	2	I38065	gene NMB protein -	914	42.5	63.4	585	2	S06958	sphingomyelin phos
842	43	64.2	560	2	I59302	brain specific Na+	915	42.5	63.4	627	1	S27393	sphingomyelin phos
843	43	64.2	574	2	T43556	Wiskott-Aldrich sy	916	42.5	63.4	629	1	A39825	sphingomyelin phos
844	43	64.2	574	2	T38819	wiskott-aldrich sy	917	42.5	63.4	679	2	B75262	conserved hypothet
845	43	64.2	575	2	C34106	protein kinase (EC	918	42.5	63.4	1047	2	A55617	masquerade precurs
846	43	64.2	591	1	WMBPQ2	gene P2 protein -	919	42.5	63.4	1211	2	T42230	AF4 protein - mous
847	43	64.2	591	2	C71460	hypothetical prote	920	42.5	63.4	1217	2	T42625	AF-4 protein - mou
848	43	64.2	592	2	T32400	hypothetical prote	921	42.5	63.4	1297	2	S25714	son-of-sevenless-2
849	43	64.2	614	2	D34106	protein kinase (EC	922	42.5	63.4	1372	2	T43296	cell fusion protei
850	43	64.2	619	2	E84800	hypothetical prote	923	42	62.7	52	2	F71353	hypothetical prote
851	43	64.2	635	1	WMBEW6	capsid protein - h	924	42	62.7	52	2	E98316	hypothetical prote
852	43	64.2	650	2	T04487	hypothetical prote	925	42	62.7	76	2	E97763	protein transport
853	43	64.2	651	2	T46050	hypothetical prote	926	42	62.7	92	2	B48831	vitelline membrane
854	43	64.2	679	2	A42073	potassium channel	927	42	62.7	97	2	E84746	hypothetical prote
855	43	64.2	691	2	F91251	probable tape meas	928	42	62.7	108	2	T26880	hypothetical prote
856	43	64.2	694	2	S68442	Grb2-associated bi	929	42	62.7	118	1	W4WL13	E4 protein - human
857	43	64.2	699	2	T05225	extensin homolog F	930	42	62.7	129	2	H75351	conserved hypothet
858	43	64.2	711	2	S43464	ecdysteroid-induce	931	42	62.7	134	2	T36365	proline-rich prote
859	43	64.2	715	2	G86239	protein F20B24.6 [932	42	62.7	135	2	T49996	AtAGP4 - Arabidops
860	43	64.2	720	2	JQ1676	ABI3 protein - Ara	933	42	62.7	143	2	S42579	QID3 protein - fun
861	43	64.2	725	2	T00492	hypothetical prote	934	42	62.7	145	2	T48552	glutaredoxin-like
862	43	64.2	727	2	C84534	hypothetical prote	935	42	62.7	149	2	T46937	hypothetical prote
863	43	64.2	733	2	C87655	penicillin-binding	936	42	62.7	159	2	D72486	hypothetical prote
864	43	64.2	769	2	I56546	Shaw type potassiu	937	42	62.7	160	2	S58759	ezrin - rat (fragm
865	43	64.2	790	2	T25095	hypothetical prote	938	42	62.7	165	2	T24470	hypothetical prote
866	43	64.2	793	1	S60735	splicing factor SF	939	42	62.7	168	2	T48343	hypothetical prote
867	43	64.2	872	2	S33015	hypothetical prote	940	42	62.7	168	2	T48375	transcription co-a
868	43	64.2	872	2	T50369	probable serine/th	941	42	62.7	173	2	T19341	hypothetical prote
869	43	64.2	876	2	T49801	SH3 domains-contai	942	42	62.7	178	2	B87104	conserved hypothet
870	43	64.2	892	2	T09071	hypothetical prote	943	42	62.7	185	2	A82752	hypothetical prote
871	43	64.2	909	2	T06635	hypothetical prote	944	42	62.7	191	2	F84522	probable proline-r
872	43	64.2	924	2	T06636	hypothetical prote	945	42	62.7	191	2	D72711	hypothetical prote
873	43	64.2	934	2	T08418	protein kinase (EC	946	42	62.7	192	2	T38535	probable transloca
874	43	64.2	971	2	T32883	hypothetical prote	947	42	62.7	194	2	AE2310	hypothetical prote
875	43	64.2	992	2	A31666	hypothetical prote	948	42	62.7	199	2	S23635	gamma-coixin, 22K,
876	43	64.2	1038	2	JT0663	ras GTPase-activat	949	42	62.7	211	2	S55129	transcription fact
877	43	64.2	1046	2	T26384	hypothetical prote	950	42	62.7	224	2	C48652	transfer protein s
878	43	64.2	1069	2	S27922	nuclear antigen EB	951	42	62.7	224	2	T46089	proline-rich prote
879	43	64.2	1086	2	T33893	hypothetical prote	952	42	62.7	265	2	T46089	En/Spm-like transp
880	43	64.2	1098	2	T08599	probable transcrip	953	42	62.7	268	2	H84684	hypothetical prote
881	43	64.2	1135	2	T30561	Scythe protein - A	954	42	62.7	287	2	T05338	hypothetical prote
882	43	64.2	1171	2	T13065	PIP82 protein - fr	955	42	62.7	293	2	G70896	hypothetical prote
883	43	64.2	1188	2	S49915	extensin-like prot	956	42	62.7	306	2	T52340	cell wall-plasma m
884	43	64.2	1201	2	A57369	anillin - fruit fl	957	42	62.7	307	2	T22128	hypothetical prote
885	43	64.2	1249	2	A56511	myosin I myoA - Em	958	42	62.7	309	2	S08343	nodulin precursor
886	43	64.2	1323	2	S27224	N-methyl-D-asparta	959	42	62.7	311	2	H70911	hypothetical prote
887	43	64.2	1323	2	I78557	N-methyl-D-asparta	960	42	62.7	318	2	S33158	extensin - common
888	43	64.2	1356	1	C45219	N-methyl-D-asparta	961	42	62.7	326	2	A46676	CD68 homolog macro
889	43	64.2	1364	2	T00250	MEGF2 protein - hu	962	42	62.7	330	2	S22140	nodulin Enod2 - Se
890	43	64.2	1371	2	T29019	hypothetical prote	963	42	62.7	342	2	A96511	unknown protein [i
891	43	64.2	1426	2	T30817	homeotic protein C	964	42	62.7	358	2	JC4311	CCAAT/enhancer bin
892	43	64.2	1557	2	T13160	protein CNK - frui	965	42	62.7	365	2	B42832	factor VIII intron
893	43	64.2	1655	2	T32633	hypothetical prote	966	42	62.7	367	2	F87340	Rieske 2Fe-2S fami
894	43	64.2	2715	2	T13049	eyelid - fruit fly	967	42	62.7	375	2	T51854	RING-H2 finger pro
895	43	64.2	2810	2	T22298	hypothetical prote	968	42	62.7	382	2	E85082	hypothetical prote
896	43	64.2	3002	2	A47221	fibrillin 1 precur	969	42	62.7	382	2	T14186	hypothetical prote
897	43	64.2	3229	2	S27852	probable cell-surf	970	42	62.7	393	2	T05532	hypothetical prote
898	43	64.2	4056	2	H96599	protein F14J16.10	971	42	62.7	394	2	T33641	hypothetical prote
899	43	64.2	4957	2	T03455	ALR protein - huma	972	42	62.7	407	1	A40150	inhibin beta-B cha
900	43	64.2	5262	2	T03454	ALR protein - huma	973	42	62.7	408	2	T34467	hypothetical prote
901	43	64.2	7962	2	I38346	elastic titin - hu	974	42	62.7	413	2	T03240	FLO/LFY protein ho
902	43	64.2	13288	2	T03099	mucin, submaxillar	975	42	62.7	414	2	JN0866	nucleolar protein
903	42.5	63.4	134	2	A95995	hypothetical prote	976	42	62.7	416	1	FOCH	transforming prote
904	42.5	63.4	150	2	C72697	hypothetical prote	977	42	62.7	416	1	TVFVAC	transforming prote
905	42.5	63.4	162	2	C72665	hypothetical prote	978	42	62.7	421	2	T00676	hypothetical prote

hypothetical prote	hypothetical prote	prpL2 protein - hu	rhodopsin - Allote	MFH-1 protein - mo	hypothetical prote	early E2A DNA-bind	hypothetical prote	sphingomyelin phos	sphingomyelin phos	conserved hypothet	masquerade precurs	AF4 protein - mous	AF-4 protein - mou	son-of-sevenless-2	cell fusion protei	hypothetical prote	hypothetical prote	protein transport	vitelline membrane	hypothetical prote	hypothetical prote	E4 protein - human	conserved hypothet	proline-rich prote	AtAGP4 - Arabidops	QID3 protein - fun	glutaredoxin-like	hypothetical prote	hypothetical prote	ezrin - rat (fragm	hypothetical prote	hypothetical prote	transcription co-a	hypothetical prote	conserved hypothet	hypothetical prote	probable proline-r	hypothetical prote	probable transloca	hypothetical prote	gamma-coixin, 22K,	transcription fact	transfer protein s	proline-rich prote	En/Spm-like transp	hypothetical prote	hypothetical prote	hypothetical prote	cell wall-plasma m	hypothetical prote	nodulin precursor	hypothetical prote	extensin - common	CD68 homolog macro	nodulin Enod2 - Se	unknown protein [i	CCAAT/enhancer bin	factor VIII intron	Rieske 2Fe-2S fami	RING-H2 finger pro	hypothetical prote	hypothetical prote	hypothetical prote	inhibin beta-B cha	hypothetical prote	FLO/LFY protein ho	nucleolar protein	transforming prote	transforming prote	hypothetical prote
--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	-------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	-------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	-------------------	--------------------	-------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	-------------------	--------------------	--------------------	--------------------

```

979 42 62.7 423 1 TVFV2E transforming prote
980 42 62.7 427 2 B96804 nucellin-like prot
981 42 62.7 429 2 S77528 hypothetical prote
982 42 62.7 440 2 S65358 familial Alzheimer
983 42 62.7 445 2 A60488 histidine-rich gly
984 42 62.7 448 2 T45710 H-protein promoter
985 42 62.7 451 1 TVFV2C gag-myc polyprotei
986 42 62.7 455 1 OOCOC rhodopsin - giant
987 42 62.7 456 2 S23104 choline kinase - h
988 42 62.7 467 2 S71169 protein kinase, 54
989 42 62.7 478 2 JC4940 synapsin IIb - hum
990 42 62.7 481 2 E86433 protein T17H7.3 [1
991 42 62.7 482 2 E97748 virB10 protein [im
992 42 62.7 483 2 A84698 probable RNA-bind
993 42 62.7 489 2 T26069 hypothetical prote
994 42 62.7 490 2 A35312 potassium channel
995 42 62.7 495 2 T27936 hypothetical prote
996 42 62.7 503 1 LUBO11 annexin XI form A
997 42 62.7 505 1 S23447 annexin XI form B
998 42 62.7 505 2 A53152 annexin XI - human
999 42 62.7 506 2 AG1864 hypothetical prote
1000 42 62.7 548 2 T14793 hypothetical prote

ALIGNMENTS

RESULT 1
I49444
SH3 binding protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I49444
R;Ren, R.; Mayer, B.J.; Cicchetti, P.; Baltimore, D.
Science 259, 1157-1161, 1993
A;Title: Identification of a ten-amino acid proline-rich SH3 binding site.
A;Reference number: I49444; MUID:93174278; PMID:8438166
A;Accession: I49444
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-559 <RES>
A;Cross-references: UNIPROT:Q06649; UNIPARC:UPI0000027A0E; GB:I14543; NID:g2932267; PIDN:
F;25-128/Domain: pleckstrin repeat homology <PLK>
F;201-210/Region: proline-rich SH3 binding

Query Match 86.6%; Score 58; DB 2; Length 559;
Best Local Similarity 80.0%; Pred. No. 8.7;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
Db 201 PPAYPPPPVP 210

RESULT 2
D41132
collagen-related protein 4 - Hydra magnipapillata (fragment)
C;Species: Hydra magnipapillata
C;Date: 05-Jun-1992 #sequence_revision 12-Jun-1992 #text_change 15-Sep-2003
C;Accession: D41132; S21932
R;Kurz, E.M.; Holstein, T.W.; Petri, B.M.; Engel, J.; David, C.N.
J. Cell Biol. 115, 1159-1169, 1991
A;Title: Mini-collagens in hydra nematocytes.
A;Reference number: A41132; MUID:92064646; PMID:1955459
A;Accession: D41132
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-172 <KUR>
A;Cross-references: UNIPARC:UPI0000076DA2; EMBL:X61048; NID:g9452; PIDN:CAA43382.1; PID:
A;Note: the authors translated the codon GGT for residue 142 as Pro, and TTA for residue
A;Note: submitted to the EMBL Data Library, July 1991

-Query Match 85.1%; Score 57; DB 2; Length 172;
```

```

Best Local Similarity 90.0%; Pred. No. 3.4;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
Db 57 PPPPPPPPIP 66

RESULT 3
S27200
proline-rich protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S27200
R;Sazuka, T.; Tomooka, Y.; Kathju, S.; Ikawa, Y.; Noda, M.; Kumar, S.
Biochim. Biophys. Acta 1132, 240-248, 1992
A;Title: Identification of a developmentally regulated gene in the mouse central nervous
A;Reference number: S27200; MUID:93041923; PMID:1420303
A;Accession: S27200
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-389 <SAZ>
A;Cross-references: UNIPROT:Q03173; UNIPARC:UPI000002954A; GB:D10727; NID:g220499; PIDN:

Query Match 85.1%; Score 57; DB 2; Length 389;
Best Local Similarity 80.0%; Pred. No. 7.8;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
Db 42 PPPLPPPPPLP 51

RESULT 4
T00645
hypothetical protein F3I6.8 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
C;Accession: T00645
R;Federspiel, N.A.; Palm, C.J.; Conway, A.B.; Kurtz, D.B.; Conway, A.R.; Au, M.; Araujo,
; Vysotskaia, V.S.; Yu, G.; Ecker, J.; Theologis, A.; Davis, R.W.
submitted to the EMBL Data Library, February 1998
A;Reference number: Z14197
A;Accession: T00645
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-820 <FED>
A;Cross-references: UNIPROT:O48682; UNIPARC:UPI00000A0F94; EMBL:AC002396; NID:g2749918;

C;Genetics:
A;Gene: ATSP:F3I6.8
A;Map position: 1
A;Introns: 335/2; 450/3; 676/2

Query Match 85.1%; Score 57; DB 2; Length 820;
Best Local Similarity 90.0%; Pred. No. 17;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
Db 238 PPPPPPPPIP 247

RESULT 5
T43500
hypothetical protein DKFZp586G1721.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
C;Accession: T43500
R;Ottenwaelder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, December 1999
A;Reference number: Z22515
A;Accession: T43500
A;Status: preliminary
```


A;Molecule type: mRNA
A;Residues: 1-222 <AAA>
A;Cross-references: UNIPARC:UPI000016AC9F; EMBL:AL133642
A;Experimental source: adult uterus; clone DKFZp586G1721
C;Genetics:
A;Note: DKFZp586G1721.1

Query Match 83.6%; Score 56; DB 2; Length 222;
Best Local Similarity 80.0%; Pred. No. 5.7;
Matches 1; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 20 PPPPPPPPPVP 29

RESULT 6
JC5614
RNB6 protein - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 23-Sep-1997 #sequence_revision 17-Oct-1997 #text_change 09-Jul-2004
C;Accession: JC5614
R;Ohta, S.; Mineta, T.; Kimoto, M.; Tabuchi, K.
Biochem. Biophys. Res. Commun. 237, 307-312, 1997
A;Title: Differential display cloning of a novel rat cDNA (RNB6) that shows high expression
A;Reference number: JC5614; MUID:97415794; PMID:9268706
A;Accession: JC5614
A;Molecule type: mRNA
A;Residues: 1-393 <OHT>
A;Cross-references: UNIPROT:O08719; UNIPARC:UPI000012A2AB; GB:U70211; NID:g2058461; PIDN
A;Experimental source: brain
C;Comment: This protein belongs to Ena/VASP family member, and is involved in the development.

Query Match 83.6%; Score 56; DB 2; Length 393;
Best Local Similarity 80.0%; Pred. No. 10;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 180 PPPPPPPPPVP 189

RESULT 7
A87378
hypothetical protein CC1037 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: A87378
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: A87378
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-563 <STO>
A;Cross-references: UNIPROT:Q9A9F1; UNIPARC:UPI00000C7254; GB:AE005673; NID:g13422331; H
C;Genetics:
A;Gene: CC1037

Query Match 83.6%; Score 56; DB 2; Length 563;
Best Local Similarity 80.0%; Pred. No. 15;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 322 PPPAPPPPLP 331

RESULT 8

Q08EA7
UL69 protein - human cytomegalovirus (strain AD169)
C;Species: human cytomegalovirus, human herpesvirus 5
A;Note: host Homo sapiens (man)
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C;Accession: S09832
R;Chee, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.; Cerny, R.; Horsnell, T.;
M.; Barrell, B.G.
Curr. Top. Microbiol. Immunol. 154, 125-169, 1990
A;Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus
A;Reference number: S09749; MUID:90269039; PMID:2161319
A;Accession: S09832
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-744 <CHE>
A;Cross-references: UNIPROT:P16749; UNIPARC:UPI000004735B; EMBL:X17403; NID:g59591; PIDN
A;Note: possible protein-coding frames are given
A;Note: the DNA sequence was submitted to EMBL, December 1989, in computer-readable form
C;Superfamily: cytomegalovirus UL69 protein
C;Keywords: transcription regulation

Query Match 83.6%; Score 56; DB 1; Length 744;
Best Local Similarity 80.0%; Pred. No. 19;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 708 PPPPPPPPPVP 717

RESULT 9
B45344
probable nuclear antigen - suid herpesvirus 1 (strain Kaplan)
C;Species: suid herpesvirus 1
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C;Accession: B45344
R;Vlcek, C.; Kozmik, Z.; Paces, V.; Schirm, S.; Schwyzer, M.
Virology 179, 365-377, 1990
A;Title: Pseudorabies virus immediate-early gene overlaps with an oppositely oriented op
A;Reference number: A45344; MUID:91021039; PMID:2171211
A;Accession: B45344
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-1733 <VLC>
A;Cross-references: UNIPROT:P33485; UNIPARC:UPI0000138ADE; GB:M34651; NID:g334070; PIDN:
C;Superfamily: pseudorabies virus 1 nuclear antigen

Query Match 83.6%; Score 56; DB 1; Length 1733;
Best Local Similarity 80.0%; Pred. No. 46;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 276 PPPRPPPPPLP 285

RESULT 10
S63244
BN11 protein - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein N0646; protein YNL271c
C;Species: Saccharomyces cerevisiae
C;Date: 27-Apr-1996 #sequence_revision 03-May-1996 #text_change 09-Jul-2004
C;Accession: S63244; S63245; S48523; S60909; S65111
R;Sen-Gupta, M.; Lyck, R.; Niedenthal, R.K.; Fleig, U.N.; Hegemann, J.H.
submitted to the Protein Sequence Database, April 1996
A;Reference number: S63235
A;Accession: S63244
A;Molecule type: DNA
A;Residues: 1-1553 <SEN>
A;Cross-references: UNIPROT:P41832; UNIPARC:UPI0000168B2A; EMBL:Z71547; MIPS:YNL271c
A;Experimental source: strain S288C
R;Messenguy, F.; Dubois, E.; Vierendeels, F.; Scherens, B.; Pierard, A.; Glansdorff, N.
submitted to the Protein Sequence Database, April 1996

A;Reference number: S63245
A;Accession: S63245
A;Molecule type: DNA
A;Residues: 987-1953 <MES>
A;Cross-references: UNIPARC:UPI000017B262; EMBL:Z71547; MIPS:YNL271c
A;Experimental source: strain S288C
R;Fares, H.F.; Pringle, J.R.
submitted to the EMBL Data Library, April 1994
A;Description: Synthetic lethals of CDC12.
A;Reference number: S48524
A;Accession: S48523
A;Molecule type: DNA
A;Residues: 1-937, 'A', 939-1429, 'C', 1431-1953 <PAR>
A;Cross-references: UNIPARC:UPI0000168E7F; EMBL:L31766; NID:g472524; PIDN:AAA34455.1; PI
R;Sen-Gupta, M.; Lyck, R.; Fleig, U.; Niedenthal, R.N.; Hegemann, J.M.
submitted to the EMBL Data Library, October 1995
A;Description: The sequence of a 24152 bp segment from the left arm of chromosome XIV fr
A;Reference number: S60909
A;Accession: S60909
A;Molecule type: DNA
A;Residues: 1-1553 <SE2>
A;Cross-references: UNIPARC:UPI0000168B2A; EMBL:X92494; NID:g1045236; PIDN:CAA63225.1; F
R;Sen-Gupta, M.; Lyck, R.; Fleig, U.; Niedenthal, R.K.; Hegemann, J.H.
Yeast 12, 505-514, 1996
A;Title: The sequence of a 24 152 bp segment from the left arm of chromosome XIV from Sa
A;Reference number: S65111; MUID:96310631; PMID:8740425
A;Accession: S65111
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1553 <SEW>
A;Cross-references: UNIPARC:UPI0000168B2A; EMBL:X92494; NID:g1045236; PIDN:CAA63225.1; F
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995
C;Genetics:
A;Gene: SGD:BNI1, SHE5, SYL39
A;Cross-references: SGD:S0005215; MIPS:YNL271c
A;Map position: 14L

Query Match 83.6%; Score 56; DB 2; Length 1953;
Best Local Similarity 80.0%; Pred. No. 51;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
||| ||||:|
Db 1241 PPPPPPPPPVP 1250

RESULT 11
B40505
hypothetical protein - suid herpesvirus 1 (strain Indiana-Funkhuser or Becker)
C;Species: suid herpesvirus 1
C;Date: 10-Apr-1992 #sequence_revision 10-Apr-1992 #text_change 09-Jul-2004
C;Accession: B40505
R;Cheung, A.K.
J. Virol. 65, 5260-5271, 1991
A;Title: Cloning of the latency gene and the early protein 0 gene of pseudorabies virus.
A;Reference number: A40505; MUID:91374576; PMID:1654441
A;Accession: B40505
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1958 <CHE>
A;Cross-references: UNIPROT:Q69340; UNIPARC:UPI00000F2308; GB:M57505; NID:g334066; PIDN:
C;Superfamily: pseudorabies virus 1 nuclear antigen

Query Match 83.6%; Score 56; DB 2; Length 1958;
Best Local Similarity 80.0%; Pred. No. 51;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
||| ||||:|
Db 485 PPPRPPPPPLP 494

RESULT 12
*

T30760
hypothetical protein 158R - Molluscum contagiosum virus 1
N;Alternate names: MC158R
C;Species: Molluscum contagiosum virus 1
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C;Accession: T30760
R;Senkevich, T.G.; Bugert, J.J.; Sisler, J.R.; Koonin, E.V.; Darai, G.; Moss, B.
Science 273, 813-816, 1996
A;Title: Genome sequence of a human tumorigenic poxvirus: Prediction of specific host re
A;Reference number: Z20876; MUID:96325459; PMID:8670425
A;Accession: T30760
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-182 <SEN>
A;Cross-references: UNIPROT:Q98324; UNIPARC:UPI00000EEB89; EMBL:U60315; PIDN:AAC55286.1
C;Genetics:
A;Note: MC158R

Query Match 82.1%; Score 55; DB 2; Length 182;
Best Local Similarity 80.0%; Pred. No. 6.1;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
||| ||||:|
Db 104 PPPPPPPPLP 113

RESULT 13
S28304
hypothetical protein T23G5.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 09-Jul-2004
C;Accession: S28304
R;Berks, M.
submitted to the EMBL Data Library, December 1992
A;Reference number: S28296
A;Accession: S28304
A;Molecule type: DNA
A;Residues: 1-211 <BER>
A;Cross-references: UNIPROT:Q03607; UNIPARC:UPI000013BA56; EMBL:Z19158; NID:g6880; PID:g
C;Genetics:
A;Introns: 138/2; 176/3

Query Match 82.1%; Score 55; DB 2; Length 211;
Best Local Similarity 80.0%; Pred. No. 7;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
||| ||||:|
Db 90 PPPPPPPPLP 99

RESULT 14
I38707
Fas ligand - human
C;Species: Homo sapiens (man)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 09-Jul-2004
C;Accession: I38707; JC2340; S57565; I38554
R;Takahashi, T.; Tanaka, M.; Inazawa, J.; Abe, T.; Suda, T.; Nagata, S.
Int. Immunol. 6, 1567-1574, 1994
A;Title: Human Fas ligand: gene structure, chromosomal location and species specificity.
A;Reference number: I38707; MUID:95127560; PMID:7826947
A;Accession: I38707
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-281 <RES>
A;Cross-references: UNIPROT:P48023; UNIPARC:UPI000000D91A; EMBL:U11821; NID:g595430; PIDN:
R;Mita, E.; Hayashi, N.; Iio, S.; Takehara, T.; Hijioke, T.; Kasahara, A.; Fusamoto, H.;
Biochem. Biophys. Res. Commun. 204, 468-474, 1994
A;Title: Role of Fas ligand in apoptosis induced by hepatitis C virus infection.
A;Reference number: JC2340; MUID:95071350; PMID:7980502
A;Accession: JC2340
A;Molecule type: DNA

A;Residues: 1-281 <MIT>
A;Cross-references: UNIPARC:UPI000000D91A; GB:D38122; DDBJ:D29820; NID:g601892; PIDN:BAZ
R;Schatzlein, C.E.
submitted to the EMBL Data Library, June 1995
A;Reference number: S57565
A;Accession: S57565
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-281 <SCH>
A;Cross-references: UNIPARC:UPI000000D91A; EMBL:X89102; NID:g887455; PID:g887456
R;Alderson, M.R.; Tough, T.W.; Davis-Smith, T.; Braddy, S.; Falk, B.; Schooley, K.A.; Go
J. Exp. Med. 181, 71-77, 1995
A;Title: Fas ligand mediates activation-induced cell death in human T lymphocytes.
A;Reference number: I38554; MUID:95105731; PMID:7528780
A;Accession: I38554
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-281 <RE2>
A;Cross-references: UNIPARC:UPI000000D91A; EMBL:U08137; NID:g624627; PIDN:AAC50071.1; PI
C;Genetics:
A;Gene: FasL
A;Introns: 151/1; 116/3
C;Keywords: glycoprotein; transmembrane protein
F;80-102/Domain: transmembrane #status predicted <TMM>
F;76,184,250,260/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 82.1%; Score 55; DB 2; Length 281;
Best Local Similarity 80.0%; Pred. No. 9.4;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
||| ||||:|
Db 46 PPPPPPPPLP 55

RESULT 15
T48514
hypothetical protein F15N18.140 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C;Accession: T48514
R;Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft,
submitted to the Protein Sequence Database, April 2000
A;Reference number: Z24490
A;Accession: T48514
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-314 <BEV>
A;Cross-references: UNIPROT:Q9LYD7; UNIPARC:UPI000000BE72; EMBL:AL163815
A;Experimental source: cultivar Columbia; BAC clone F15N18
C;Genetics:
A;Map position: 5
A;Note: F15N18.140

Query Match 82.1%; Score 55; DB 2; Length 314;
Best Local Similarity 80.0%; Pred. No. 10;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
|||:|||||
Db 81 PPPHPPPPPP 90

RESULT 16
S02193
cellular tumor antigen p53 - chicken
N;Alternate names: nuclear oncoprotein p53
C;Species: Gallus gallus (chicken)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S02193
R;Soussi, T.; Begue, A.; Kress, M.; Stehelin, D.; May, P.
Nucleic Acids Res. 16, 11383, 1988
A;Title: Nucleotide sequence of a cDNA encoding the chicken p53 nuclear oncoprotein.

A;Reference number: S02193; MUID:89083584; PMID:3060861
A;Accession: S02193
A;Molecule type: mRNA
A;Residues: 1-367 <SOU>
A;Cross-references: UNIPROT:P10360; UNIPARC:UPI000013103A; EMBL:X13057; NID:g63740; PIDN
C;Superfamily: cellular tumor antigen p53
C;Keywords: apoptosis; cell division control; DNA binding; homotetramer; nucleus; phosph
F;161,164,224,228/Binding site: zinc (Cys, His, Cys, Cys) #status predicted
F;366/Binding site: phosphoryl-RNA (Ser) (covalent) #status predicted

Query Match 82.1%; Score 55; DB 1; Length 367;
Best Local Similarity 80.0%; Pred. No. 12;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
||| ||||:|
Db 50 PPPPPPPPLP 59

RESULT 17
T07910
hydroxyproline-rich glycoprotein GAS29 precursor - Chlamydomonas reinhardtii
C;Species: Chlamydomonas reinhardtii
C;Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 09-Jul-2004
C;Accession: T07910
R;Rodriguez-Martinez, H.; Haring, M.A.; Von Gromof, E.; Beck, C.F.
submitted to the EMBL Data Library, July 1997
A;Reference number: Z16207
A;Accession: T07910
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-433 <ROD>
A;Cross-references: UNIPROT:O22459; UNIPARC:UPI00000A1D78; EMBL:AF015884; NID:g2384729;
A;Experimental source: gametes
C;Genetics:
A;Gene: GAS29
F;1-31/Domain: signal sequence #status predicted <SIG>
F;32-433/Product: hydroxyproline-rich glycoprotein GAS29 #status predicted <MAT>

Query Match 82.1%; Score 55; DB 2; Length 433;
Best Local Similarity 80.0%; Pred. No. 15;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
||| ||||:|
Db 63 PPPPPPPPLP 72

RESULT 18
T34018
hypothetical protein Y4C6B.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T34018
R;Layman, D.; Graves, T.; Yoakum, M.
submitted to the EMBL Data Library, February 1999
A;Description: The sequence of C. elegans cosmid Y4C6B.
A;Reference number: Z21460
A;Accession: T34018
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-442 <LAY>
A;Cross-references: UNIPROT:Q9UB01; UNIPARC:UPI0000074EDD; EMBL:AF125971; PIDN:AAD14764.
A;Experimental source: strain Bristol N2; clone Y4C6B
C;Genetics:
A;Gene: CESP:Y4C6B.1
A;Map position: 4
A;Introns: 43/3; 78/3; 268/3; 387/2
C;Superfamily: Caenorhabditis elegans hypothetical protein Y4C6B.1

Query Match 82.1%; Score 55; DB 2; Length 442;
Best Local Similarity 80.0%; Pred. No. 15;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
||| |||||:|
Db 315 PPPPPPPPLP 324

RESULT 19
B85187
glycoprotein homolog [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C/Accession: B85187
R/anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Nature 402, 769-777, 1999
A/Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A/Reference number: A85001; MUID:20083488; PMID:10617198
A/Accession: B85187
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-473 <STO>
A/Cross-references: UNIPROT:Q9SUK6; UNIPARC:UPI00000AC27B; GB:NC_001268; NID:g7268429; R/Genetics:
A/Gene: At4g16790
A/Map position: 4

Query Match 82.1%; Score 55; DB 2; Length 473;
Best Local Similarity 80.0%; Pred. No. 16;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
||| |||||:|
Db 284 PPPPPPPPLP 293

RESULT 20
E97813
WASP, N-WASP, MENA proteins homolog [imported] - Rickettsia conorii (strain Malish 7)
C/Species: Rickettsia conorii
C/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C/Accession: E97813
R/Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; R/Science 293, 2093-2098, 2001
A/Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A/Reference number: A97700; MUID:21442074; PMID:11557893
A/Accession: E97813
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-520 <KUR>
A/Cross-references: UNIPROT:Q92H62; UNIPARC:UPI00000D4EF6; GB:AE006914; PIDN:AAL03447.1;
C/Genetics:
A/Gene: RC0909

Query Match 82.1%; Score 55; DB 2; Length 520;
Best Local Similarity 80.0%; Pred. No. 17;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
||| |||||:|
Db 352 PPPPPPPPLP 361

RESULT 21
I38607
p53-binding protein 2 - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 09-Mar-1996 #sequence_revision 09-Mar-1996 #text_change 31-Dec-2004
C/Accession: I38607
R/Iwabuchi, K.; Bartel, P.L.; Li, B.; Marraccino, R.; Fields, S./Proc. Natl. Acad. Sci. U.S.A. 91, 6098-6102, 1994
A/Title: Two cellular proteins that bind to wild-type but not mutant p53.
A/Reference number: I38604; MUID:94286584; PMID:8016121
A/Accession: I38607
A/Status: preliminary; nucleic acid sequence not shown

A/Molecule type: mRNA
A/Residues: 1-529 <RES>
A/Cross-references: UNIPROT:Q13625; UNIPARC:UPI000016A0A8; EMBL:U09582; NID:g493079; PID:F/392-424/Domain: ankyrin repeat homology <AN08>
F/465-515/Domain: SH3 homology <SH3>

Query Match 82.1%; Score 55; DB 2; Length 529;
Best Local Similarity 88.9%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PPYPppppPIP 10
||| |||||:|
Db 268 PPYPppppYP 276

RESULT 22
D40228
neurexin II-beta precursor - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 21-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
C/Accession: D40228; S27888
R/Ushkaryov, Y.A.; Petrenko, A.G.; Geppert, M.; Suedhof, T.C./Science 257, 50-56, 1992
A/Title: Neurexins: synaptic cell surface proteins related to the alpha-latrotoxin receptor
A/Reference number: A40228; MUID:92320296; PMID:1621094
A/Accession: D40228
A/Status: preliminary; nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1-662 <USH>
A/Cross-references: UNIPROT:Q63376; UNIPARC:UPI0000130AA5; GB:M96377; NID:g205717; PIDN:A/Note: authors translated GAC for residue 411 as Thr and ACC for residue 412 as Asp
C/Superfamily: neurexin; EGF homology
C/Keywords: alternative splicing; transmembrane protein
F/1-47/Domain: signal sequence #status predicted <SIG>
F/48-662/Product: neurexin II-beta #status predicted <MAT>

Query Match 82.1%; Score 55; DB 2; Length 662;
Best Local Similarity 80.0%; Pred. No. 22;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPYPppppPIP 10
||| |||||:|
Db 23 PPPPPPPPLP 32

RESULT 23
JC5929
serine/arginine-rich protein-specific kinase (EC 2.-.-.-) 2 - mouse
C/Species: Mus musculus (house mouse)
C/Date: 10-Apr-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C/Accession: JC5929
R/Kuroyanagi, N.; Onogi, H.; Wakabayashi, T.; Hagiwara, M./Biochem. Biophys. Res. Commun. 242, 357-364, 1998
A/Title: Novel SR-protein-specific kinase, SRPK2, disassembles nuclear speckles.
A/Reference number: JC5929; MUID:98113357; PMID:9446799
A/Accession: JC5929
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1-681 <KUR>
A/Cross-references: UNIPROT:Q8VCD9; UNIPROT:O54781; UNIPARC:UPI000017C748
A/Experimental source: brain
C/Comment: This enzyme regulates the dissembly of the serine/arginine-rich proteins in a
C/Keywords: transferase

Query Match 82.1%; Score 55; DB 2; Length 681;
Best Local Similarity 80.0%; Pred. No. 23;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPYPppppPIP 10
||| |||||:|
Db 27 PPPPPPPPLP 36

RESULT 24
T46476
hypothetical protein DKFZp434C0931.1 - human
C;Species: Homo sapiens (man)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C;Accession: T46476
R;Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23034
A;Accession: T46476
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-691 <AAA>
A;Cross-references: UNIPARC:UPI000015FC7C; EMBL:AL137718
A;Experimental source: adult testis; clone DKFZp434C0931
C;Genetics:
A;Note: DKFZp434C0931.1

Query Match 82.1%; Score 55; DB 2; Length 691;
Best Local Similarity 80.0%; Pred. No. 23;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
||| |||||:
Db 328 PPPPPPPPLP 337

RESULT 25
D96563
probable bZIP protein, 48652-45869 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: D96563
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
anssen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: D96563
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-798 <STO>
A;Cross-references: UNIPROT:Q9C824; UNIPARC:UPI00000A2A93; GB:AE005173; NID:g10645444; F
C;Genetics:
A;Gene: F19K6.7
A;Map position: 1

Query Match 82.1%; Score 55; DB 2; Length 798;
Best Local Similarity 80.0%; Pred. No. 27;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
||| |||||:
Db 116 PPPPPPPPLP 125

RESULT 26
E96636
hypothetical protein T7P1.21 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: E96636
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
anssen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: E96636
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-907 <STO>
A;Cross-references: UNIPROT:Q9C946; UNIPARC:UPI00000A9620; GB:AE005173; NID:g6751696; PI
C;Genetics:
A;Gene: T7P1.21
A;Map position: 1

Query Match 82.1%; Score 55; DB 2; Length 907;
Best Local Similarity 80.0%; Pred. No. 31;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
||| |||||:
Db 508 PPPPPPPPLP 517

RESULT 27
S19595
chloride channel protein - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S19595
R;Steinmeyer, K.; Ortland, C.; Jentsch, T.J.
Nature 354, 301-304, 1991
A;Title: Primary structure and functional expression of a developmentally regulated skel
A;Reference number: S19595; MUID:92065954; PMID:1659664
A;Accession: S19595
A;Molecule type: mRNA
A;Residues: 1-994 <STE>
A;Cross-references: UNIPROT:P35524; UNIPARC:UPI0000127A8F; GB:X62894; NID:g57744; PIDN:C
C;Keywords: transmembrane protein
F;830-877/Domain: CBS homology <CBS>

Query Match 82.1%; Score 55; DB 2; Length 994;
Best Local Similarity 80.0%; Pred. No. 34;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
||| |||||:
Db 728 PPPPPPPPLP 737

RESULT 28
T13286
cappuccino gene protein - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C;Accession: T13286
R;Emmons, S.; Phan, H.; Calley, J.; Chen, W.; James, B.; Manseau, L.
Genes Dev. 9, 2482-2494, 1995
A;Title: Cappuccino, a Drosophila maternal effect gene required for polarity of the egg
A;Reference number: Z17651; MUID:96033799; PMID:7590229
A;Accession: T13286
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1058 <ENM>
A;Cross-references: UNIPROT:Q24120; UNIPARC:UPI000016BA5C; EMBL:U34258; NID:g1061333; PI
C;Genetics:
A;Gene: capu
A;Cross-references: FlyBase:FBgn00000256

Query Match 82.1%; Score 55; DB 2; Length 1058;
Best Local Similarity 80.0%; Pred. No. 36;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
||| ||||:|
Db 486 PPPPPPPPLP 495

RESULT 29
Tl3963
formin related protein, lymphocyte specific - mouse
C;Species: Mus musculus (house mouse)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: Tl3963
R;Taniuchi, I.; Yamamoto, S.; Watanabe, T.
submitted to the EMBL Data Library, June 1997
A;Description: Lymphocyte specific formin related gene.
A;Reference number: Z17840
A;Accession: Tl3963
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1064 <TAN>
A;Cross-references: UNIPROT:Q9Z2V7; UNIPARC:UPI000002989C; EMBL:AF006466; NID:g4101719;
C;Genetics:
A;Gene: Fr1

Query Match 82.1%; Score 55; DB 2; Length 1064;
Best Local Similarity 80.0%; Pred. No. 36;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
||| ||||:|
Db 514 PPPPPPPPLP 523

RESULT 30
Tl7454
diaphanous-related formin - mouse
C;Species: Mus musculus (house mouse)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: Tl7454
R;Tomimaga, T.; Sahai, E.; Treisman, R.H.; Alberts, A.S.
submitted to the EMBL Data Library, September 1998
A;Reference number: Z18796
A;Accession: Tl7454
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1171 <TOM>
A;Cross-references: UNIPROT:Q9Z207; UNIPARC:UPI0000029711; EMBL:AF094519; NID:g3845724;
C;Genetics:
A;Gene: Dia2

Query Match 82.1%; Score 55; DB 2; Length 1171;
Best Local Similarity 80.0%; Pred. No. 40;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
||| ||||:|
Db 571 PPPPPPPPLP 580

RESULT 31
G86441
unknown protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: G86441
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maity, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: G86441
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1201 <STO>
A;Cross-references: UNIPROT:Q9C6S1; UNIPARC:UPI000009C9CB; GB:AE005172; NID:g11136725; P
C;Genetics:
A;Map position: 1

Query Match 82.1%; Score 55; DB 2; Length 1201;
Best Local Similarity 80.0%; Pred. No. 41;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
||| ||||:|
Db 574 PPPPPPPPLP 583

RESULT 32
S24407
formin isoform IV - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S24407
R;Jackson-Grusby, L.; Kuo, A.; Leder, P.
Genes Dev. 6, 29-37, 1992
A;Title: A variant limb deformity transcript expressed in the embryonic mouse limb define
A;Reference number: S24407; MUID:92112033; PMID:1339380
A;Accession: S24407
A;Molecule type: mRNA
A;Residues: 1-1206 <JAC>
A;Cross-references: UNIPROT:Q05859; UNIPARC:UPI0000027920; EMBL:X62379; NID:g51552; PIDN

Query Match 82.1%; Score 55; DB 2; Length 1206;
Best Local Similarity 80.0%; Pred. No. 41;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
||| ||||:|
Db 705 PPPPPPPPLP 714

RESULT 33
T31065
diaphanous protein homolog p140mDia - mouse
C;Species: Mus musculus (house mouse)
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C;Accession: T31065
R;Watanabe, N.; Madaule, P.; Reid, T.; Ishizaki, T.; Watanabe, G.; Kakizuka, A.; Saito, Y
EMBO J. 16, 3044-3056, 1997
A;Title: p140mDia, a mammalian homolog of Drosophila diaphanous, is a target protein for
A;Reference number: Z20961; MUID:97357293; PMID:9214622
A;Accession: T31065
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1255 <WAT>
A;Cross-references: UNIPROT:O08808; UNIPARC:UPI0000027B7C; EMBL:U96963; NID:g2114472; PII
A;Note: binds to GTP-bound form of Rho and binds to profilin

Query Match 82.1%; Score 55; DB 2; Length 1255;
Best Local Similarity 80.0%; Pred. No. 43;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
||| ||||:|
Db 586 PPPPPPPPLP 595

RESULT 34
S48375
hypothetical protein YII159w - yeast (Saccharomyces cerevisiae)
C;Species: Saccharomyces cerevisiae

C;Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 09-Jul-2004

C;Accession: S48375
R;Churcher, C.
submitted to the EMBL Data Library, September 1994
A;Reference number: S48310
A;Accession: S48375
A;Molecule type: DNA
A;Residues: 1-1375 <CHU>
A;Cross-references: UNIPROT:P40450; UNIPARC:UPI0000126A45; GB:Z47047; EMBL:Z38059; NID:9
C;Genetics:
A;Gene: SGD:BNR1
A;Cross-references: SGD:S0001421; MIPS:YIL159W
A;Map position: 9L

Query Match 82.1%; Score 55; DB 2; Length 1375;
Best Local Similarity 80.0%; Pred. No. 47;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPIP 10

||| |||||:
Db 772 PPPPPPPPLP 781

RESULT 35

S11515
formin - mouse
C;Species: Mus musculus (house mouse)
C;Date: 22-Jan-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S11515
R;Woychik, R.P.; Maas, R.L.; Zeller, R.; Vogt, T.F.; Leder, P.
Nature 346, 850-853, 1990
A;Title: 'Formins': proteins deduced from the alternative transcripts of the limb deform
A;Reference number: S11515; MUID:90363291; PMID:2392150
A;Accession: S11515
A;Molecule type: mRNA
A;Residues: 1-1468 <WOY>
A;Cross-references: UNIPROT:Q05860; UNIPARC:UPI0000027927; EMBL:X53599; NID:G52877; PIDN

Query Match 82.1%; Score 55; DB 2; Length 1468;
Best Local Similarity 80.0%; Pred. No. 50;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPIP 10

||| |||||:
Db 931 PPPPPPPPLP 940

RESULT 36

F97713
190K antigen precursor [imported] - Rickettsia conorii (strain Malish 7)
C;Species: Rickettsia conorii
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: F97713
R;Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Ro
Science 293, 2093-2098, 2001
A;Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A;Reference number: A97700; MUID:21442074; PMID:11557893
A;Accession: F97713
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1795 <KUR>
A;Cross-references: UNIPROT:Q92JF7; UNIPARC:UPI00000CBCAD; GB:AE006914; PIDN:AAL02648.1;
C;Genetics:
A;Gene: sca2

Query Match 82.1%; Score 55; DB 2; Length 1795;
Best Local Similarity 80.0%; Pred. No. 61;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPIP 10

||| |||||:
Db 1078 PPPPPPPPLP 1087

RESULT 37

A41948
alpha-fetoprotein enhancer-binding protein - human
N;Alternate names: ATBF1 protein
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 15-Oct-1999
C;Accession: A41948
R;Moringa, T.; Yasuda, H.; Hashimoto, T.; Higashio, K.; Tamaoki, T.
Mol. Cell. Biol. 11, 6041-6049, 1991
A;Title: A human alpha-fetoprotein enhancer-binding protein, ATBF1, contains four homeod
A;Reference number: A41948; MUID:92049333; PMID:1719379
A;Accession: A41948
A;Molecule type: mRNA
A;Residues: 1-2783 <MOR>
A;Cross-references: UNIPARC:UPI0000156381; GB:D10250; GB:D90395; NID:G219429; PIDN:BAA01
A;Note: sequence extracted from NCBI backbone (NCBIN:66271, NCBIIP:66276)
C;Genetics:
A;Gene: GDB:ATBF1
A;Cross-references: GDB:392090; OMIM:104155
A;Map position: 16q22.3-16q23.1
C;Superfamily: alpha-fetoprotein enhancer-binding protein; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation; zinc finger
F;72-94/Region: zinc finger CCHH motif
F;128-150/Region: zinc finger CCHH motif
F;176-198/Region: zinc finger CCHH motif
F;311-332/Region: zinc finger CCHH motif
F;340-361/Region: zinc finger CCHH motif
F;448-471/Region: zinc finger CCHH motif
F;489-509/Region: zinc finger CCHH motif
F;517-538/Region: zinc finger CCHH motif
F;633-655/Region: zinc finger CCHH motif
F;684-706/Region: zinc finger CCHH motif
F;719-773/Region: serine/threonine-rich
F;809-958/Region: glutamine-rich
F;1071-1092/Region: zinc finger CCHH motif
F;1117-1211/Region: proline-rich
F;1232-1288/Domain: homeobox homology <HOX1>
F;1329-1385/Domain: homeobox homology <HOX2>
F;1416-1437/Region: zinc finger CCHH motif
F;1618-1638/Region: zinc finger CCHH motif
F;1728-1784/Domain: homeobox homology <HOX3>
F;1799-1820/Region: zinc finger CCHH motif
F;2033-2089/Domain: homeobox homology <HOX4>
F;2112-2134/Region: zinc finger CCHH motif
F;2545-2566/Region: zinc finger CCHH motif
F;2585-2607/Region: glycine-rich
F;2611-2633/Region: zinc finger CCHH motif
F;2650-2737/Region: serine/threonine-rich

Query Match 82.1%; Score 55; DB 1; Length 2783;
Best Local Similarity 80.0%; Pred. No. 95;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPIP 10

||| |||||:
Db 1129 PPPPPPPPLP 1138

RESULT 38

T28957
hypothetical protein F45F2.11 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T28957
R;Davidson, S.; Wohldmann, P.
submitted to the EMBL Data Library, July 1996
A;Description: The sequence of C. elegans cosmid F45F2.
A;Reference number: Z20548
A;Accession: T28957
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-406 <DAV>

A;Cross-references: UNIPROT:Q22975; UNIPARC:UPI000007865A; EMBL:U64845; PIDN:AAC48032.1;
A;Experimental source: strain Bristol N2; clone F45F2
C;Genetics:
A;Gene: CESP:F45F2.11
A;Map position: 5
A;Introns: 1/2; 24/2; 143/3; 307/2; 338/2

Query Match 80.6%; Score 54; DB 2; Length 406;
Best Local Similarity 77.8%; Pred. No. 18;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 PPYP PPPPIP 10
||:|||||:
Db 109 PFP PPPPLP 117

RESULT 39
T31502
hypothetical protein Y116A8C.32 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T31502
R;McMurray, A.
submitted to the EMBL Data Library, October 1999
A;Reference number: Z21041
A;Accession: T31502
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-710 <WIL>
A;Cross-references: UNIPROT:Q9U2U1; UNIPARC:UPI000017BC48; EMBL:AL117204; PIDN:CAB55136.
A;Experimental source: clone Y116A8C
C;Genetics:
A;Gene: CESP:Y116A8C.32
A;Introns: 53/1; 188/2; 238/3; 257/2; 308/2; 437/2; 600/1

Query Match 80.6%; Score 54; DB 2; Length 710;
Best Local Similarity 80.0%; Pred. No. 31;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPYP PPPPIP 10
|||:||||:
Db 685 PPP PPPPMP 694

RESULT 40
T13170
diaphanous protein - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jul-2004
C;Accession: T13170
R;Castrillon, D.H.; Wasserman, S.A.; Castrillon, D.H.; Wasserman, S.A.
Development 120, 3367-3377, 1994
A;Title: Diaphanous is required for cytokinesis in Drosophila and shares domains of simi
A;Reference number: Z17626; MUID:95121197; PMID:7821209
A;Accession: T13170
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1091 <CAS>
A;Cross-references: UNIPROT:P48608; UNIPARC:UPI000016BA0C; EMBL:U11288; NID:g575926; PID
C;Genetics:
A;Gene: dia
A;Cross-references: FlyBase:FBgn0011202
A;Map position: 2L

Query Match 80.6%; Score 54; DB 2; Length 1091;
Best Local Similarity 80.0%; Pred. No. 48;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPYP PPPPIP 10
|||:||||:
Db 539 PPP PPPPMP 548

RESULT 41
D48232
cysteine-rich extensin-like protein 4 precursor - common tobacco (fragment)
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 09-Jul-2004
C;Accession: D48232
R;Wu, H.; Zou, J.; May, B.; Gu, Q.; Cheung, A.Y.
Proc. Natl. Acad. Sci. U.S.A. 90, 6829-6833, 1993
A;Title: A tobacco gene family for flower cell wall proteins with a proline-rich domain
A;Reference number: A48232; MUID:93342083; PMID:8341705
A;Accession: D48232
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-160 <WUA>
A;Cross-references: UNIPROT:Q08197; UNIPARC:UPI0000177E6F; GB:113442
A;Note: authors failed to translate the codon GCT for residue 11 as Ala, and GTA for res
A;Note: authors translated the codon TCT for residue 60 as Pro, and CCA for residue 117
C;Genetics:
A;Gene: CELP-4
C;Superfamily: glutelin
C;Keywords: cell wall; extracellular matrix; fertilization; glycoprotein
F;1-17/Domain: signal sequence #status predicted <SIG>
F;18-159/Product: cysteine-rich extensin-like protein 4 #status experimental <MAT>

Query Match 79.1%; Score 53; DB 2; Length 160;
Best Local Similarity 80.0%; Pred. No. 8.9;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PPYP PPPPIP 10
|||:||||:
Db 74 PPC PPPPAP 83

RESULT 42
E48232
cysteine-rich extensin-like protein 5 precursor - common tobacco
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 09-Jul-2004
C;Accession: E48232
R;Wu, H.; Zou, J.; May, B.; Gu, Q.; Cheung, A.Y.
Proc. Natl. Acad. Sci. U.S.A. 90, 6829-6833, 1993
A;Title: A tobacco gene family for flower cell wall proteins with a proline-rich domain
A;Reference number: A48232; MUID:93342083; PMID:8341705
A;Accession: E48232
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-161 <WUA>
A;Cross-references: UNIPROT:Q08198; UNIPARC:UPI00000A99C3; GB:113443; NID:g310930; PIDN:
C;Genetics:
A;Gene: CELP-5
C;Superfamily: glutelin
C;Keywords: cell wall; extracellular matrix; fertilization; glycoprotein
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-160/Product: cysteine-rich extensin-like protein 5 #status experimental <MAT>

Query Match 79.1%; Score 53; DB 2; Length 161;
Best Local Similarity 80.0%; Pred. No. 9;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PPYP PPPPIP 10
|||:||||:
Db 75 PPC PPPPAP 84

RESULT 43
C48232
cysteine-rich extensin-like protein 3 precursor - common tobacco
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 09-Jul-2004
C;Accession: C48232
R;Wu, H.; Zou, J.; May, B.; Gu, Q.; Cheung, A.Y.
Proc. Natl. Acad. Sci. U.S.A. 90, 6829-6833, 1993
A;Title: A tobacco gene family for flower cell wall proteins with a proline-rich domain

A;Reference number: A48232; MUID:93342083; PMID:8341705
A;Accession: C48232
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-165 <WUA>
A;Cross-references: UNIPROT:Q08196; UNIPARC:UPI00000AA5D7; GB:L13441; NID:g310926; PIDN:
A;Note: authors translated the codon GAT for residue 35 as Glu
C;Genetics:
A;Gene: CELP-3
C;Superfamily: glutelin
C;Keywords: cell wall; extracellular matrix; fertilization; glycoprotein
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-165/Product: cysteine-rich extensin-like protein 3 #status experimental <MAT>

Query Match 79.1%; Score 53; DB 2; Length 165;
Best Local Similarity 80.0%; Pred. No. 9.2;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PPPYPPPPIP 10
||| |||||
Db 77 PPPSPPPAP 86

RESULT 44
B48232
cysteine-rich extensin-like protein 2 precursor - common tobacco
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 26-May-1994 #sequence revision 26-May-1994 #text_change 09-Jul-2004
C;Accession: B48232; PQ0474; S24616
R;Wu, H.; Zou, J.; May, B.; Gu, Q.; Cheung, A.Y.
Proc. Natl. Acad. Sci. U.S.A. 90, 6829-6833, 1993
A;Title: A tobacco gene family for flower cell wall proteins with a proline-rich domain
A;Reference number: A48232; MUID:93342083; PMID:8341705
A;Accession: B48232
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-196 <WUA>
A;Cross-references: UNIPROT:Q08195; UNIPARC:UPI00000A3E44; GB:L13440; NID:g310924; PIDN:
R;de S Goldman, M.H.; Pezzotti, M.; Seurinck, J.; Mariani, C.
Plant Cell 4, 1041-1051, 1992
A;Title: Developmental expression of tobacco pistil-specific genes encoding novel extens
A;Reference number: PQ0474; MUID:93005740; PMID:1392607
A;Accession: PQ0474
A;Molecule type: mRNA
A;Residues: 'MAG', 1-105 <GOL>
A;Cross-references: UNIPARC:UPI0000177E6E; EMBL:Z14014
A;Experimental source: stigma, style; strain Petit Havana SR1
C;Genetics:
A;Gene: CELP-2
C;Superfamily: glutelin
C;Keywords: cell wall; extracellular matrix; fertilization
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-196/Product: cysteine-rich extensin-like protein 2 #status experimental <MAT>

Query Match 79.1%; Score 53; DB 2; Length 196;
Best Local Similarity 80.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PPPYPPPPIP 10
||| |||||
Db 107 PPPSPPPAP 116

RESULT 45
A48232
cysteine-rich extensin-like protein 1 precursor - common tobacco
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 26-May-1994 #sequence revision 26-May-1994 #text_change 09-Jul-2004
C;Accession: A48232; PQ0475; S24617
R;Wu, H.; Zou, J.; May, B.; Gu, Q.; Cheung, A.Y.
Proc. Natl. Acad. Sci. U.S.A. 90, 6829-6833, 1993
A;Title: A tobacco gene family for flower cell wall proteins with a proline-rich domain
A;Reference number: A48232; MUID:93342083; PMID:8341705

A;Accession: A48232
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-209 <WUA>
A;Cross-references: UNIPROT:Q08194; UNIPARC:UPI00000ABB3F; GB:L13439; NID:g310922; PIDN:
R;de S Goldman, M.H.; Pezzotti, M.; Seurinck, J.; Mariani, C.
Plant Cell 4, 1041-1051, 1992
A;Title: Developmental expression of tobacco pistil-specific genes encoding novel extens
A;Reference number: PQ0474; MUID:93005740; PMID:1392607
A;Accession: PQ0475
A;Molecule type: mRNA
A;Residues: 39-209 <GOL>
A;Cross-references: UNIPARC:UPI000016DE8A; EMBL:Z14020; NID:g19918; PID:g19919
A;Experimental source: stigma, style; strain Petit Havana SR1
C;Superfamily: glutelin
C;Keywords: cell wall; extracellular matrix; fertilization; glycoprotein
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-209/Product: cysteine-rich extensin-like protein 1 #status experimental <MAT>
F;146/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 79.1%; Score 53; DB 2; Length 209;
Best Local Similarity 80.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PPPYPPPPIP 10
||| |||||
Db 119 PPPSPPPAP 128

RESULT 46
E86255
hypothetical protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: E86255
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: E86255
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-744 <STO>
A;Cross-references: UNIPROT:O65375; UNIPARC:UPI00000A65A5; GB:AE005172; NID:g3157926; PI
C;Genetics:
A;Map position: 1

Query Match 79.1%; Score 53; DB 2; Length 744;
Best Local Similarity 80.0%; Pred. No. 42;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PPPYPPPPIP 10
||| |||||
Db 525 PPPSPPPPCP 534

RESULT 47
T32404
hypothetical protein R148.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T32404
R;Le, T.T.; Kemp, K.; Scheet, P.
submitted to the EMBL Data Library, September 1997
A;Description: The sequence of C. elegans cosmid R148.
A;Reference number: Z21161

A;Accession: T32404
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1127 <LET>
A;Cross-references: UNIPROT:O17267; UNIPARC:UPI0000176698; EMBL:AF025467; PIDN:AAB71038.
A;Experimental source: strain Bristol N2; clone R148
C;Genetics:
A;Gene: CESP:R148.3
A;Map position: 3
A;Introns: 11/3; 42/1; 87/3; 373/3; 544/2; 595/2; 1042/1
C;Superfamily: Caenorhabditis elegans hypothetical protein R148.3

Query Match 79.1%; Score 53; DB 2; Length 1127;
Best Local Similarity 80.0%; Pred. No. 64;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
||| |||||
Db 388 PPPAPPPPPAP 397

RESULT 48
JC5765
inositol polyphosphate 5-phosphatase (EC 3.1.3.-) 2 - human
C;Species: Homo sapiens (man)
C;Date: 24-Jan-1998 #sequence_revision 13-Mar-1998 #text_change 09-Jul-2004
C;Accession: JC5765
R;Pesesse, X.; Deleu, S.; De Smedt, F.; Drayer, L.; Erneux, C.
Biochem. Biophys. Res. Commun. 239, 697-700, 1997
A;Title: Identification of a second SH2-domain-containing protein closely related to the
A;Reference number: JC5765; MUID:98042455; PMID:9367831
A;Accession: JC5765
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-1258 <PES>
A;Cross-references: UNIPROT:O15357; UNIPARC:UPI000003AEB7; GB:Y14385; NID:g2653423; PIDN
C;Comment: This enzyme plays a role in metabolism of inositol phosphate and phosphatidyl
C;Keywords: phosphoric monoester hydrolase
F;21-117/Domain: SH2 homology <SH2>
F;427-729/Domain: catalytic #status predicted <CAT>
F;927-1171/Region: proline-rich

Query Match 79.1%; Score 53; DB 2; Length 1258;
Best Local Similarity 70.0%; Pred. No. 72;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
|| :|||:|
Db 1050 PPDFPPPPPLP 1059

RESULT 49
T00381
KIAA0633 protein - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
C;Accession: T00381
R;Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.
DNA Res. 5, 169-176, 1998
A;Title: Prediction of the coding sequences of unidentified human genes. X. The complete
A;Reference number: Z14142; MUID:98403880; PMID:9734811
A;Accession: T00381
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1316 <ISH>
A;Cross-references: UNIPROT:O75128; UNIPARC:UPI0000070BAF; EMBL:AB014533; NID:g3327079;
C;Genetics:
A;Gene: KIAA0633

Query Match 79.1%; Score 53; DB 2; Length 1316;
Best Local Similarity 80.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
||| |||||
Db 394 PPPQPPPPSP 403

RESULT 50
PQ0452
extensin-like protein - Persian tobacco (strain S2S3) (fragment)
C;Species: Nicotiana alata (Persian tobacco)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 05-Oct-2004
C;Accession: PQ0452
R;Chen, C.G.; Cornish, B.C.; Clarke, A.E.
Plant Cell 4, 1053-1062, 1992
A;Title: Specific expression of an extensin-like gene in the style of Nicotiana alata.
A;Reference number: PQ0451; MUID:93005741; PMID:1392608
A;Accession: PQ0452
A;Molecule type: mRNA
A;Residues: 1-122 <CHE>
A;Cross-references: UNIPROT:Q41192; UNIPARC:UPI0000177CC3
A;Experimental source: style, strain S2S3

Query Match 77.6%; Score 52; DB 2; Length 122;
Best Local Similarity 80.0%; Pred. No. 8.8;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
||| |||||
Db 54 PPPSPPPPPSP 63

Search completed: April 6, 2006, 09:39:52
Job time : 32.7895 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2006, 09:23:24 ; Search time 145.895 Seconds
(without alignments)
58.030 Million cell updates/sec

Title: US-10-632-388-312
Perfect score: 67
Sequence: 1 PPPYPPPIPPX 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : UniProt_05.80.*

- 1: uniprot_sprot.*
- 2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query	Length	DB ID	Description
1	59	88.1	88.1	480	2	Q740X0 MYCPA
2	58	86.6	86.6	347	2	Q6Z1Z7 ORYSA
3	58	86.6	86.6	467	2	Q6ZTK4 HUMAN
4	58	86.6	86.6	536	2	Q6ZVU3 HUMAN
5	58	86.6	86.6	539	2	Q5ZHV9 CHICK
6	58	86.6	86.6	559	1	3BP2 MOUSE
7	58	86.6	86.6	559	2	Q5U3L0 MOUSE
8	58	86.6	86.6	559	2	Q91Z52 MOUSE
9	58	86.6	86.6	559	2	Q8C220 MOUSE
10	58	86.6	86.6	561	1	3BP2 HUMAN
11	58	86.6	86.6	1166	2	Q7PWW3 ANOGA
12	58	86.6	86.6	1428	2	Q6Z1Z8 ORYSA
13	57	85.1	85.1	172	2	Q00487 9CNID
14	57	85.1	85.1	222	2	Q6AVF5 ORYSA
15	57	85.1	85.1	231	2	Q5AX03 EMENI
16	57	85.1	85.1	282	1	ATF5 HUMAN
17	57	85.1	85.1	434	2	Q8IWI2 HUMAN
18	57	85.1	85.1	442	2	Q6VUC0 HUMAN
19	57	85.1	85.1	521	2	Q61WU8 CAEBR
20	57	85.1	85.1	590	2	Q5SIV4 THET8
21	57	85.1	85.1	802	1	ENAH MOUSE
22	57	85.1	85.1	820	2	Q48682 ARATH
23	57	85.1	85.1	955	1	FRU DROME
24	57	85.1	85.1	1134	2	Q96JH1 HUMAN
25	57	85.1	85.1	1146	2	Q6PLP6 CHLRE
26	57	85.1	85.1	1273	2	Q4WGB0 ASPFU
27	57	85.1	85.1	1324	2	Q8MMQ2 DICDI
28	57	85.1	85.1	1395	2	Q7SC01 NEUCR
29	57	85.1	85.1	1400	2	Q9VDD2 DROME
30	57	85.1	85.1	1918	2	Q752A6 ASHGO
31	56	83.6	83.6	135	2	Q5JM22 ORYSA

105	55	82.1	243	2	Q75JU4_DICDI	Q75JU4 dictyosteli	178	55	82.1	526	2	Q5XHX3_RAT	Q5xhx3 rattus norv
106	55	82.1	247	2	Q5W6V8_ORYSA	Q5w6v8 oryza sativ	179	55	82.1	532	2	Q9QY28_MOUSE	Q9qy28 mus musculu
107	55	82.1	262	1	YN03_CAEL	Q03607 caenorhabdi	180	55	82.1	533	2	Q80ZC2_MOUSE	Q80zc2 mus musculu
108	55	82.1	268	2	Q9DGN3_CHICK	Q9dgn3 gallus gall	181	55	82.1	541	2	Q5D053_MOUSE	Q5d053 mus musculu
109	55	82.1	280	1	TNFL6_CERTO	Q9bdn1 c tumor nec	182	55	82.1	550	2	Q93263_CHICK	Q93263 gallus gall
110	55	82.1	280	1	TNFL6_MACFA	P63308 m tumor nec	183	55	82.1	558	2	Q6PIM5_HUMAN	Q6pim5 homo sapien
111	55	82.1	280	1	TNFL6_MACMU	P63307 m tumor nec	184	55	82.1	558	2	Q5BJ88_XENTR	Q5bj88 xenopus tro
112	55	82.1	280	1	TNFL6_MACNE	P63306 m tumor nec	185	55	82.1	559	1	WASF1_MOUSE	Q8r5h6 mus musculu
113	55	82.1	280	2	Q861W5_FELCA	Q861w5 felis silve	186	55	82.1	559	1	WASF1_MOUSE	Q8r5h6 mus musculu
114	55	82.1	281	1	TNFL6_HUMAN	P48023 h tumor nec	187	55	82.1	559	2	Q5SZK7_HUMAN	Q5szk7 homo sapien
115	55	82.1	281	2	Q53ZZ1_HUMAN	Q53zz1 homo sapien	188	55	82.1	559	2	Q5NVG8_PONPY	Q5nv8 pongo pygma
116	55	82.1	283	2	Q4RDJ1_TETNG	Q4rdj1 tetraodon n	189	55	82.1	559	2	Q5BJU7_RAT	Q5bj7 rattus norv
117	55	82.1	285	2	Q9LS54_ARATH	Q9ls54 arabidopsis	190	55	82.1	560	2	Q6GLN2_XENLA	Q6gln2 xenopus lae
118	55	82.1	285	2	Q6GYS3_RAT	Q6gys3 rattus norv	191	55	82.1	568	2	Q96MU3_HUMAN	Q96mu3 homo sapien
119	55	82.1	285	2	Q6GYS1_MOUSE	Q6gys1 mus musculu	192	55	82.1	579	2	Q4RLQ7_TETNG	Q4rlq7 tetraodon n
120	55	82.1	288	2	Q6XYB2_HUMAN	Q6xyb2 homo sapien	193	55	82.1	591	1	ENAH_HUMAN	Q8n8r7 homo sapien
121	55	82.1	289	2	Q8IXZ1_HUMAN	Q8ixz1 homo sapien	194	55	82.1	591	2	Q502W5_HUMAN	Q502w5 homo sapien
122	55	82.1	289	2	Q5BL00_HUMAN	Q5bl00 homo sapien	195	55	82.1	605	2	Q9Q5L3_9GAMA	Q9q5l3 cercopithec
123	55	82.1	290	1	HXD8_HUMAN	P13378 homo sapien	196	55	82.1	608	2	Q69S58_ORYSA	Q69s58 oryza sativ
124	55	82.1	296	2	Q5CWS3_CRYPV	Q5cws3 cryptospori	197	55	82.1	631	2	Q8K2Q9_MOUSE	Q8k2q9 mus musculu
125	55	82.1	306	2	Q085S1_MOUSE	Q085s1 mus musculu	198	55	82.1	634	2	Q8N9B5_HUMAN	Q8n9b5 homo sapien
126	55	82.1	307	2	Q4RSI9_TETNG	Q4rsi9 tetraodon n	199	55	82.1	649	2	Q68DG1_HUMAN	Q68dgl1 homo sapien
127	55	82.1	314	2	Q9LYD7_ARATH	Q9lyd7 arabidopsis	200	55	82.1	662	1	NRX2B_RAT	Q63376 rattus norv
128	55	82.1	315	2	Q6ASR7_ORYSA	Q6asr7 oryza sativ	201	55	82.1	669	2	Q6MZF0_HUMAN	Q6mzf0 homo sapien
129	55	82.1	315	2	Q6GYS2_RAT	Q6gys2 rattus norv	202	55	82.1	671	2	Q9S5S9_HUMAN	Q9s5s9 homo sapien
130	55	82.1	315	2	Q6GYS0_MOUSE	Q6gys0 mus musculu	203	55	82.1	673	2	Q4S708_TETNG	Q4s708 tetraodon n
131	55	82.1	321	2	Q9PTD5_CHICK	Q9ptd5 gallus gall	204	55	82.1	681	1	SRPK2_MOUSE	Q54781 mus musculu
132	55	82.1	326	2	Q6UUS0_ORYSA	Q6uus0 oryza sativ	205	55	82.1	681	2	Q9VD05_DROME	Q9vdo5 drosophila
133	55	82.1	328	1	MTFR1_MOUSE	Q99mb2 mus musculu	206	55	82.1	686	2	Q5DTW5_MOUSE	Q5dtw5 mus musculu
134	55	82.1	328	2	Q8CGA8_MOUSE	Q8cga8 mus musculu	207	55	82.1	687	2	Q4R9A2_MACFA	Q4r9a2 macaca fasc
135	55	82.1	333	2	Q99PM7_MOUSE	Q99pm7 mus musculu	208	55	82.1	688	1	SPRK2_HUMAN	P78362 homo sapien
136	55	82.1	359	2	Q7FAV6_ORYSA	Q7fav6 oryza sativ	209	55	82.1	692	2	Q581R1_9TRYP	Q581r1 trypanosoma
137	55	82.1	359	2	Q7FAP4_ORYSA	Q7fap4 oryza sativ	210	55	82.1	692	2	Q4SUB2_TETNG	Q4sub2 tetraodon n
138	55	82.1	359	2	Q7XP48_ORYSA	Q7xp48 oryza sativ	211	55	82.1	693	2	Q4T7R3_TETNG	Q4t7r3 tetraodon n
139	55	82.1	360	2	Q96HE9_HUMAN	Q96he9 homo sapien	212	55	82.1	694	2	Q6F359_ORYSA	Q6f359 oryza sativ
140	55	82.1	360	2	Q9NUZ7_HUMAN	Q9nuz7 homo sapien	213	55	82.1	699	2	Q8NAV4_HUMAN	Q8nav4 homo sapien
141	55	82.1	367	1	P53_CHICK	P10360 gallus gall	214	55	82.1	723	2	Q5T2S7_HUMAN	Q5t2s7 homo sapien
142	55	82.1	368	2	Q8BHE0_MOUSE	Q8bhe0 m riken cdn	215	55	82.1	724	2	Q80VK9_MOUSE	Q80vk9 mus musculu
143	55	82.1	420	2	Q9DEG1_CHICK	Q9deg1 gallus gall	216	55	82.1	735	2	Q4T2R8_TETNG	Q4t2r8 tetraodon n
144	55	82.1	427	2	Q6ETL2_ORYSA	Q6etl2 oryza sativ	217	55	82.1	737	2	Q5RE41_PONPY	Q5re41 pongo pygma
145	55	82.1	433	2	Q22459_CHLRE	Q22459 chlamydomon	218	55	82.1	737	2	Q8H4U8_ORYSA	Q8h4u8 oryza sativ
146	55	82.1	435	2	Q9SM17_MAIZE	Q9sm17 zea mays (m	219	55	82.1	742	2	Q9FW12_ORYSA	Q9fw12 oryza sativ
147	55	82.1	435	2	Q80ZB7_RAT	Q80zb7 rattus norv	220	55	82.1	749	2	Q8NB24_HUMAN	Q8nb24 homo sapien
148	55	82.1	442	2	Q9UB01_CAEL	Q9ub01 caenorhabdi	221	55	82.1	771	2	Q4R7E4_MACFA	Q4r7e4 macaca fasc
149	55	82.1	443	2	Q80ZC1_MOUSE	Q80zc1 mus musculu	222	55	82.1	784	2	Q53SQ1_HUMAN	Q53sq1 homo sapien
150	55	82.1	444	2	Q80ZB6_RAT	Q80zb6 rattus norv	223	55	82.1	784	2	Q90YB5_CHICK	Q90yb5 gallus gall
151	55	82.1	445	2	Q95KJ7_MACFA	Q95kj7 macaca fasc	224	55	82.1	795	2	Q504V9_HUMAN	Q504v9 homo sapien
152	55	82.1	446	2	Q9HCH4_HUMAN	Q9hch4 homo sapien	225	55	82.1	798	2	Q9C824_ARATH	Q9c824 arabidopsis
153	55	82.1	447	2	Q4SQR1_TETNG	Q4sqrl1 tetraodon n	226	55	82.1	798	2	Q571B6_MOUSE	Q571b6 mus musculu
154	55	82.1	451	2	Q62775_RAT	Q62775 rattus norv	227	55	82.1	824	2	Q8C414_MOUSE	Q8c414 mus musculu
155	55	82.1	452	2	Q80ZC0_MOUSE	Q80zc0 mus musculu	228	55	82.1	825	2	Q7S9H3_NEUCR	Q7s9h3 neurospora
156	55	82.1	456	2	Q8C4F6_MOUSE	Q8c4f6 mus musculu	229	55	82.1	830	2	Q6NRA4_XENLA	Q6nra4 xenopus lae
157	55	82.1	464	2	Q9FF14_ARATH	Q9ff14 arabidopsis	230	55	82.1	835	2	Q6PGC9_MOUSE	Q6pgc9 mus musculu
158	55	82.1	470	2	Q9LEM1_9BRAS	Q9lem1 capsella ru	231	55	82.1	836	2	Q4WU48_ASPFU	Q4wu48 aspergillus
159	55	82.1	471	2	Q9M3R6_ARATH	Q9m3r6 arabidopsis	232	55	82.1	837	2	Q63618_RAT	Q63618 rattus norv
160	55	82.1	473	2	Q8LCM1_ARATH	Q8lcm1 arabidopsis	233	55	82.1	844	2	Q6ZSY1_HUMAN	Q6zsy1 homo sapien
161	55	82.1	473	2	Q9SUK6_ARATH	Q9suk6 arabidopsis	234	55	82.1	849	2	Q6NUP0_HUMAN	Q6nup0 homo sapien
162	55	82.1	483	2	Q6ZK43_ORYSA	Q6zk43 oryza sativ	235	55	82.1	849	2	Q5XKF6_HUMAN	Q5xkf6 homo sapien
163	55	82.1	485	2	Q9Z0G8_RAT	Q9z0g8 rattus norv	236	55	82.1	853	1	DIAP3_HUMAN	Q9nav4 homo sapien
164	55	82.1	489	2	Q7YTX4_DROME	Q7ytx4 drosophila	237	55	82.1	853	2	Q4SJT0_TETNG	Q4sjt0 tetraodon n
165	55	82.1	494	2	Q96HR8_HUMAN	Q96hr8 homo sapien	238	55	82.1	854	2	Q5JXM8_HUMAN	Q5jxm8 homo sapien
166	55	82.1	494	2	Q9AKJ0_RICRI	Q9akj0 rickettsia	239	55	82.1	854	2	Q9H0A2_HUMAN	Q9h0a2 homo sapien
167	55	82.1	495	2	Q4RWN6_TETNG	Q4rwn6 tetraodon n	240	55	82.1	854	2	Q5T8W7_HUMAN	Q5t8w7 homo sapien
168	55	82.1	497	2	Q5RA03_PONPY	Q5ra03 pongo pygma	241	55	82.1	871	2	Q9ET47_MOUSE	Q9et47 mus musculu
169	55	82.1	501	2	Q4T4F3_TETNG	Q4t4f3 tetraodon n	242	55	82.1	884	2	Q9MA60_ARATH	Q9ma60 arabidopsis
170	55	82.1	502	2	Q7XRX7_ORYSA	Q7xrx7 oryza sativ	243	55	82.1	907	1	ANDR_CANFA	Q9tt90 canis famil
171	55	82.1	511	2	Q80ZB9_RAT	Q80zb9 rattus norv	244	55	82.1	907	2	Q9C946_ARATH	Q9c946 arabidopsis
172	55	82.1	513	2	Q9DEG2_CHICK	Q9deg2 gallus gall	245	55	82.1	922	2	Q7TT75_RAT	Q7tt75 rattus norv
173	55	82.1	517	2	Q8LN36_ORYSA	Q8ln36 oryza sativ	246	55	82.1	927	2	Q52ZP1_HUMAN	Q52zpl1 homo sapien
174	55	82.1	520	2	Q92H62_RICCN	Q92h62 rickettsia	247	55	82.1	929	2	Q8K331_MOUSE	Q8k331 mus musculu
175	55	82.1	520	2	Q80ZB8_RAT	Q80zb8 rattus norv	248	55	82.1	964	2	Q7TT74_RAT	Q7tt74 rattus norv
176	55	82.1	521	2	Q612S1_CAEBR	Q612s1 caenorhabdi	249	55	82.1	971	1	R3HDM_HUMAN	Q15032 homo sapien
477	55	82.1	524	2	Q80ZC3_MOUSE	Q80zc3 mus musculu	250	55	82.1	980	2	Q60PF6_CAEBR	Q60pf6 caenorhabdi

251	55	82.1	983	2	Q9QXM1_MOUSE	Q9qxm1 mus musculus
252	55	82.1	983	2	Q5BL16_MOUSE	Q5bl16 mus musculus
253	55	82.1	994	1	CLCN1_RAT	P35524 rattus norv
254	55	82.1	996	2	Q5CKJ5_CRYHO	Q5ckj5 cryptospori
255	55	82.1	998	2	Q6GNM2_XENLA	Q6gnm2 xenopus lae
256	55	82.1	1009	2	Q8N3E0_HUMAN	Q8n3e0 homo sapien
257	55	82.1	1037	2	Q73808_FUGRU	Q73808 fugu rubrip
258	55	82.1	1051	2	Q9SE97_ARATH	Q9se97 arabidopsis
259	55	82.1	1060	2	Q6NSN9_BRARE	Q6nsn9 brachydanio
260	55	82.1	1061	2	Q5R7K5_PONPY	Q5r7k5 pongo pygma
261	55	82.1	1087	2	Q59L24_CANAL	Q59l24 candida alb
262	55	82.1	1088	1	ASPP2_MOUSE	Q8cg79 mus musculus
263	55	82.1	1094	1	FMNL_MOUSE	Q9jl26 mus musculus
264	55	82.1	1098	1	DIAP2_MOUSE	Q70566 mus musculus
265	55	82.1	1099	2	Q8IW32_HUMAN	Q8iw32 homo sapien
266	55	82.1	1102	2	Q6W4W7_MOUSE	Q6w4w7 mus musculus
267	55	82.1	1107	2	Q54Z23_DICDI	Q54z23 dictyosteli
268	55	82.1	1112	2	Q96PY5_HUMAN	Q96py5 homo sapien
269	55	82.1	1125	2	Q66IP5_XENLA	Q66ip5 xenopus lae
270	55	82.1	1128	1	ASPP2_HUMAN	Q13625 homo sapien
271	55	82.1	1130	2	Q9C0D6_HUMAN	Q9c0d6 homo sapien
272	55	82.1	1139	2	Q6NTK5_XENLA	Q6ntk5 xenopus lae
273	55	82.1	1141	2	Q4KLM6_XENLA	Q4klw6 xenopus lae
274	55	82.1	1145	2	Q4KLV4_XENLA	Q4klv4 xenopus lae
275	55	82.1	1147	2	Q5PR81_MOUSE	Q5pr81 mus musculus
276	55	82.1	1156	2	Q7XGQ4_ORYSA	Q7xgg4 oryza sativ
277	55	82.1	1156	2	Q8W2Q5_ORYSA	Q8w2q5 oryza sativ
278	55	82.1	1165	2	Q4T2J8_TETNG	Q4t2j8 tetraodon n
279	55	82.1	1167	2	Q640S7_XENTR	Q640s7 xenopus tro
280	55	82.1	1171	1	DIAP3_MOUSE	Q9z207 mus musculus
281	55	82.1	1175	2	Q4W6X2_COPCI	Q4w6x2 coprinus ci
282	55	82.1	1192	2	Q5DTM7_MOUSE	Q5dtm7 mus musculus
283	55	82.1	1201	2	Q9C6S1_ARATH	Q9c6s1 arabidopsis
284	55	82.1	1204	2	Q4S986_TETNG	Q4s986 tetraodon n
285	55	82.1	1206	1	FMN1B_MOUSE	Q05859 mus musculus
286	55	82.1	1216	2	Q4SL06_TETNG	Q4sl06 tetraodon n
287	55	82.1	1220	2	Q6NS79_MOUSE	Q6ns79 mus musculus
288	55	82.1	1236	2	Q9C0I2_HUMAN	Q9c0i2 homo sapien
289	55	82.1	1248	1	DIAP1_HUMAN	Q06010 homo sapien
290	55	82.1	1250	2	Q70E73_HUMAN	Q70e73 homo sapien
291	55	82.1	1255	1	DIAP1_MOUSE	Q08808 mus musculus
292	55	82.1	1269	2	Q7XH56_ORYSA	Q7xh56 oryza sativ
293	55	82.1	1269	2	Q8W5K6_ORYSA	Q8w5k6 oryza sativ
294	55	82.1	1271	2	Q5KSL6_HUMAN	Q5ksl6 homo sapien
295	55	82.1	1272	2	Q6URC4_HUMAN	Q6urc4 homo sapien
296	55	82.1	1278	2	Q4RQM1_TETNG	Q4rqm1 tetraodon n
297	55	82.1	1283	2	Q6CKG2_KLULA	Q6ckg2 kluyveromyc
298	55	82.1	1285	2	Q5DTQ4_MOUSE	Q5dtq4 mus musculus
299	55	82.1	1286	2	Q5BHU8_DROME	Q5bhu8 drosophila
300	55	82.1	1289	2	Q9FLQ7_ARATH	Q9flq7 arabidopsis
301	55	82.1	1294	2	Q6FRQ4_CANGA	Q6frq4 candida gla
302	55	82.1	1299	2	Q59FH8_HUMAN	Q59fh8 homo sapien
303	55	82.1	1326	2	Q5XGY6_XENLA	Q5xgy6 xenopus lae
304	55	82.1	1347	2	Q4RW30_TETNG	Q4rw30 tetraodon n
305	55	82.1	1375	1	BNR1_YEAST	P40450 saccharomyc
306	55	82.1	1454	2	Q5SPA0_BRARE	Q5spa0 brachydanio
307	55	82.1	1466	2	Q56B22_RICAF	Q56b22 rickettsia
308	55	82.1	1468	1	FMN1A_MOUSE	Q05860 mus musculus
309	55	82.1	1527	1	ARHGB_RAT	Q9es67 rattus norv
310	55	82.1	1535	2	Q66JL8_MOUSE	Q66jl8 mus musculus
311	55	82.1	1552	2	Q68FM7_MOUSE	Q68fm7 mus musculus
312	55	82.1	1627	2	Q84ZL0_ORYSA	Q84zl0 oryza sativ
313	55	82.1	1635	2	Q7X693_ORYSA	Q7x693 oryza sativ
314	55	82.1	1641	2	Q5NCY0_MOUSE	Q5ncy0 mus musculus
315	55	82.1	1641	2	Q4VC26_MOUSE	Q4vc26 mus musculus
316	55	82.1	1652	2	Q9C0F0_HUMAN	Q9c0f0 homo sapien
317	55	82.1	1682	2	Q15054_HUMAN	Q15054 homo sapien
318	55	82.1	1707	2	Q6P9K2_MOUSE	Q6p9k2 mus musculus
319	55	82.1	1732	2	Q5AL52_CANAL	Q5al52 candida alb
320	55	82.1	1739	1	DOT1L_HUMAN	Q8tek3 homo sapien
321	55	82.1	1755	2	Q6BPE0_DEBHA	Q6bpe0 debaryomyce
322	55	82.1	1755	2	Q4IC42_GIBZE	Q4ic42 gibberella
323	55	82.1	1764	2	Q51VS4_MAGGR	Q51vs4 magnaporthe

324	55	82.1	1766	2	Q8NF45_HUMAN	Q8nf45 homo sapien
325	55	82.1	1795	2	Q92JF7_RICCN	Q92jf7 rickettsia
326	55	82.1	1835	2	Q6BNW7_DEBHA	Q6bnw7 debaryomyce
327	55	82.1	1842	2	Q6CUT9_KLULA	Q6cut9 kluyveromyc
328	55	82.1	1851	2	Q6C9I8_YARLI	Q6c9i8 yarrowia li
329	55	82.1	1883	2	Q8LMR2_ORYSA	Q8lmr2 oryza sativ
330	55	82.1	1884	2	Q4RWK6_TETNG	Q4rwk6 tetraodon n
331	55	82.1	1951	1	YLPM1_HUMAN	P49750 homo sapien
332	55	82.1	2068	2	Q67WF0_ORYSA	Q67wf0 oryza sativ
333	55	82.1	2147	2	Q61N21_CAENR	Q61n21 caenorhabdi
334	55	82.1	2333	2	Q615U0_CAEBR	Q615u0 caenorhabdi
335	55	82.1	3567	2	Q86UP3_HUMAN	Q86up3 homo sapien
336	55	82.1	3703	1	ATBF1_HUMAN	Q15911 homo sapien
337	55	82.1	3726	1	ATBF1_MOUSE	Q61329 mus musculu
338	55	82.1	4969	2	Q8CF9I_MOUSE	Q8cf9i mus musculu
339	55	82.1	5038	1	PCLO_MOUSE	Q9gyx7 mus musculu
340	55	82.1	5120	1	PCLO_CHICK	Q9pu36 gallus gall
341	55	82.1	5165	2	Q8CF92_MOUSE	Q8cf92 mus musculu
342	55	82.1	5183	1	PCLO_HUMAN	Q9y6v0 homo sapien
343	54.5	81.3	222	2	Q5TN00_ANOGA	Q5tn00 anopheles g
344	54.5	81.3	3070	2	Q4RLJ2_TETNG	Q4rlj2 tetraodon n
345	54	80.6	101	2	Q8BR19_MOUSE	Q8br19 mus musculu
346	54	80.6	129	2	Q5ZE72_ORYSA	Q5ze72 oryza sativ
347	54	80.6	192	2	Q7S944_NEUCR	Q7s944 neurospora
348	54	80.6	213	2	Q4KT55_9NUCL	Q4kt55 chrysodeixi
349	54	80.6	219	2	Q61UQ5_CAEBR	Q61uq5 caenorhabdi
350	54	80.6	352	2	Q6EP36_ORYSA	Q6ep36 oryza sativ
351	54	80.6	358	2	O82239_ARATH	O82239 arabidopsis
352	54	80.6	389	2	Q6ET49_ORYSA	Q6et49 oryza sativ
353	54	80.6	401	2	Q9U3U1_CAEBR	Q9u3u1 caenorhabdi
354	54	80.6	406	2	Q22975_CAEL	Q22975 caenorhabdi
355	54	80.6	414	2	Q624L9_CAEBR	Q624l9 caenorhabdi
356	54	80.6	417	2	Q4KT78_9NUCL	Q4kt78 chrysodeixi
357	54	80.6	462	2	Q9IBR4_9NUCL	Q9ibr4 spodoptera
358	54	80.6	485	1	PDCD7_HUMAN	Q8n8d1 homo sapien
359	54	80.6	485	2	Q6IEG3_HUMAN	Q6ieg3 homo sapien
360	54	80.6	515	2	Q7YTY2_DROME	Q7yty2 drosophila
361	54	80.6	515	2	Q9VCH2_DROME	Q9vch2 drosophila
362	54	80.6	522	2	Q60XF4_CAEBR	Q60xf4 caenorhabdi
363	54	80.6	664	2	Q628E1_CAEBR	Q628e1 caenorhabdi
364	54	80.6	699	2	Q9U2U1_CAEL	Q9u2u1 caenorhabdi
365	54	80.6	705	2	Q8IVU2_HUMAN	Q8ivu2 homo sapien
366	54	80.6	722	2	Q6ZVG1_HUMAN	Q6zvg1 homo sapien
367	54	80.6	732	2	Q8J1Y5_ASHGO	Q8jly5 ashbya goss
368	54	80.6	732	2	Q74ZB4_ASHGO	Q74zb4 ashbya goss
369	54	80.6	788	2	Q69MT2_ORYSA	Q69mt2 oryza sativ
370	54	80.6	789	2	Q4SHS7_TETNG	Q4shs7 tetraodon n
371	54	80.6	1005	2	Q6ZVH7_HUMAN	Q6zvh7 homo sapien
372	54	80.6	1091	1	DIA_DROME	P48608 drosophila
373	54	80.6	1091	2	Q5BI26_DROME	Q5bi26 drosophila
374	54	80.6	1217	2	Q9ULL5_HUMAN	Q9ull5 homo sapien
375	54	80.6	1408	2	Q7Q976_ANOGA	Q7q976 anopheles g
376	54	80.6	1485	2	Q4I763_GIBZE	Q4i763 gibberella
377	54	80.6	1533	2	Q7SAT8_NEUCR	Q7sat8 neurospora
378	54	80.6	1612	2	Q7PV65_ANOGA	Q7pv65 anopheles g
379	54	80.6	1790	2	Q86ZG7_NEUCR	Q86zg7 neurospora
380	54	80.6	1817	2	Q7RWH7_NEUCR	Q7rwh7 neurospora
381	53	79.1	132	2	Q61H28_CAEBR	Q61h28 caenorhabdi
382	53	79.1	157	2	Q08197_TOBAC	Q08197 nicotiana t
383	53	79.1	161	2	Q08198_TOBAC	Q08198 nicotiana t
384	53	79.1	165	2	Q08196_TOBAC	Q08196 nicotiana t
385	53	79.1	171	2	Q07DLN2_TOBAC	Q07dln2 nicotiana t
386	53	79.1	196	2	Q08195_TOBAC	Q08195 nicotiana t
387	53	79.1	209	2	Q08194_TOBAC	Q08194 nicotiana t
388	53	79.1	223	2	Q5VR17_ORYSA	Q5vr17 oryza sativ
389	53	79.1	235	2	Q02442_GLORO	Q02442 globodera r
390	53	79.1	239	2	Q9VRU9_DROME	Q9vrU9 drosophila
391	53	79.1	246	2	Q5VRF4_ORYSA	Q5vrf4 oryza sativ
392	53	79.1	273	2	Q80YB9_MOUSE	Q80yb9 mus musculu
393	53	79.1	337	2	Q8IVB5_HUMAN	Q8ivb5 homo sapien
394	53	79.1	338	2	Q5PQO7_RAT	Q5pgq7 rattus norv
395	53	79.1	347	2	Q585A8_9TRYP	Q585a8 trypanosoma
396	53	79.1	347	2	Q585B0_9TRYP	Q585b0 trypanosoma

397	53	79.1	348	2	Q8BQ89_MOUSE	Q8bq89 mus musculus	470	52	77.6	243	2	Q67UD2_ORYSA	Q67ud2 oryza sativ
398	53	79.1	360	2	Q9XZM6_STRPU	Q9xzm6 strongyloce	471	52	77.6	247	2	Q5AY44_EMENI	Q5ay44 aspergillus
399	53	79.1	379	2	Q504Y4_HUMAN	Q504y4 homo sapien	472	52	77.6	251	2	Q58WM7_9POAL	Q58wm7 elionurus t
400	53	79.1	400	2	Q9SQF7_BRAJU	Q9sqf7 brassica ju	473	52	77.6	251	2	Q8L3S3_ORYSA	Q8l3s3 oryza sativ
401	53	79.1	415	2	Q8S9B6_VOLCA	Q8s9b6 volvox cart	474	52	77.6	256	2	Q5YB71_MANSE	Q5yb71 manduca sex
402	53	79.1	479	2	Q08611_MOUSE	Q08611 mus musculus	475	52	77.6	256	2	Q58WL1_9POAL	Q58wl1 ischaemum a
403	53	79.1	491	2	Q7RXL8_NEUCR	Q7rxl8 neurospora	476	52	77.6	258	2	Q58WM3_9POAL	Q58wm3 coelorachis
404	53	79.1	520	2	Q5VR46_ORYSA	Q5vr46 oryza sativ	477	52	77.6	258	2	Q58WM4_9POAL	Q58wm4 apluda muti
405	53	79.1	744	2	Q65375_ARATH	Q65375 arabidopsis	478	52	77.6	258	2	Q58WM5_9POAL	Q58wm5 coelorachis
406	53	79.1	899	2	Q8N304_HUMAN	Q8n304 homo sapien	479	52	77.6	260	2	Q58WK7_COILA	Q58wk7 coix lachry
407	53	79.1	1017	2	Q4SYS7_TETNG	Q4sys7 tetraodon n	480	52	77.6	260	2	Q58WM2_9POAL	Q58wm2 coix aquati
408	53	79.1	1068	2	Q6CCCO_YARLI	Q6ccco yarrowia li	481	52	77.6	264	2	Q6VEQ4_HUMAN	Q6veq4 homo sapien
409	53	79.1	1087	2	Q7Q1T1_ANOGA	Q7qlt1 anopheles g	482	52	77.6	264	2	Q6VEQ5_HUMAN	Q6veq5 homo sapien
410	53	79.1	1115	2	Q17267_CAEEL	Q17267 caenorhabdi	483	52	77.6	264	2	Q6P447_HUMAN	Q6p447 homo sapien
411	53	79.1	1130	2	Q8ITY8_CAEEL	Q8ity8 caenorhabdi	484	52	77.6	264	2	Q4NY85_9DELT	Q4ny85 anaeromyxob
412	53	79.1	1146	2	Q86XA7_HUMAN	Q86xa7 homo sapien	485	52	77.6	268	1	LBD13_ARATH	Q9at61 arabidopsis
413	53	79.1	1149	2	Q13577_HUMAN	Q13577 homo sapien	486	52	77.6	270	2	Q7Z666_HUMAN	Q7z666 homo sapien
414	53	79.1	1175	2	Q6A0A9_MOUSE	Q6a0a9 mus musculus	487	52	77.6	271	2	Q8NFB6_HUMAN	Q8nfb6 homo sapien
415	53	79.1	1183	2	Q9R1V2_RAT	Q9rlv2 rattus norv	488	52	77.6	275	2	Q7Z7K6_HUMAN	Q7z7k6 homo sapien
416	53	79.1	1257	2	Q9JLL7_MOUSE	Q9jll7 mus musculus	489	52	77.6	297	2	Q6ZT55_HUMAN	Q6zt55 homo sapien
417	53	79.1	1257	2	Q9WVR3_RAT	Q9wvr3 rattus norv	490	52	77.6	297	2	Q67K16_SYMTH	Q67k16 symbiobacte
418	53	79.1	1257	2	Q6P549_MOUSE	Q6p549 mus musculus	491	52	77.6	299	2	Q7X838_ORYSA	Q7x838 oryza sativ
419	53	79.1	1258	2	Q15357_HUMAN	Q15357 homo sapien	492	52	77.6	300	2	Q4THG6_TETNG	Q4thh6 tetraodon n
420	53	79.1	1316	2	Q75128_HUMAN	Q75128 homo sapien	493	52	77.6	301	2	Q6L4D1_ORYSA	Q6l4d1 oryza sativ
421	53	79.1	1391	2	Q6K8Z4_ORYSA	Q6k8z4 oryza sativ	494	52	77.6	302	2	Q5Z7H9_ORYSA	Q5z7h9 oryza sativ
422	53	79.1	4027	2	Q5I2R0_CHLIN	Q5i2r0 chlamydomon	495	52	77.6	302	2	Q9CAC7_ARATH	Q9cac7 arabidopsis
423	52	77.6	39	2	Q9BR93_HUMAN	Q9br93 homo sapien	496	52	77.6	305	2	Q9SH44_ARATH	Q9sh44 arabidopsis
424	52	77.6	56	2	Q8MM17_ACRPR	Q8mm17 acropora pr	497	52	77.6	309	2	Q6IEL3_ORYSA	Q6iel3 oryza sativ
425	52	77.6	56	2	Q8MM18_ACRPL	Q8mm18 acropora pa	498	52	77.6	309	2	Q7XNY3_ORYSA	Q7xny3 oryza sativ
426	52	77.6	56	2	Q8MM19_ACRCE	Q8mm19 acropora ce	499	52	77.6	314	2	Q54817_MOUSE	Q54817 mus musculus
427	52	77.6	56	2	Q8MUZ3_ACRPL	Q8muz3 acropora pa	500	52	77.6	318	2	Q5PQY8_BRARE	Q5pqy8 brachydanio
428	52	77.6	56	2	Q8MUZ4_ACRPL	Q8muz4 acropora pa	501	52	77.6	323	1	LECT_SOLTU	Q9s8m0 solanum tub
429	52	77.6	56	2	Q8MUZ5_9CNID	Q8muz5 acropora na	502	52	77.6	329	2	Q9WX60_ACEXY	Q9wx60 acetobacter
430	52	77.6	74	2	Q59EF8_HUMAN	Q59ef8 homo sapien	503	52	77.6	333	2	Q7XR83_ORYSA	Q7xr83 oryza sativ
431	52	77.6	83	2	Q9GTI9_HETGL	Q9gti9 heterodera	504	52	77.6	344	2	Q9H7P6_HUMAN	Q9h7p6 homo sapien
432	52	77.6	87	2	Q7WXH7_ALCEU	Q7wxh7 alcaligenes	505	52	77.6	351	2	Q7QD64_ANOGA	Q7qd64 anopheles g
433	52	77.6	90	2	Q4VMJ0_MANSE	Q4vmj0 manduca sex	506	52	77.6	353	2	Q6TRW9_CHICK	Q6trw9 gallus gall
434	52	77.6	115	2	Q5AL62_CANAL	Q5al62 candida alb	507	52	77.6	356	2	Q6TRX0_CHICK	Q6trx0 gallus gall
435	52	77.6	116	2	Q8C8A6_MOUSE	Q8c8a6 mus musculus	508	52	77.6	359	2	Q9XZT0_DROME	Q9xzt0 drosophila
436	52	77.6	137	2	Q4JMU1_9BACT	Q4jmul uncultured	509	52	77.6	361	2	Q6IMV8_ORYSA	Q6imv8 oryza sativ
437	52	77.6	138	2	Q7QE51_ANOGA	Q7qe51 anopheles g	510	52	77.6	365	2	Q4KLX2_XENLA	Q4klx2 xenopus lae
438	52	77.6	138	2	Q9SSK0_ARATH	Q9ssk0 arabidopsis	511	52	77.6	367	2	Q59847_ASPOR	Q59847 aspergillus
439	52	77.6	144	2	Q23120_ARATH	Q23120 arabidopsis	512	52	77.6	373	2	Q62CV6_BURMA	Q62cv6 burkholderi
440	52	77.6	144	2	Q8LAZ0_ARATH	Q8laz0 arabidopsis	513	52	77.6	376	2	Q53N23_ORYSA	Q53n23 oryza sativ
441	52	77.6	145	2	Q8MLF7_DROME	Q8mlf7 drosophila	514	52	77.6	378	2	Q19961_CAEEL	Q19961 caenorhabdi
442	52	77.6	147	2	Q8CBK2_MOUSE	Q8cbk2 mus musculus	515	52	77.6	378	2	Q6T9Z5_CHICK	Q6t9z5 gallus gall
443	52	77.6	149	2	Q00484_9CNID	Q00484 hydra sp. m	516	52	77.6	382	2	Q88933_MOUSE	Q88933 mus musculus
444	52	77.6	149	2	Q35327_MOUSE	Q35327 mus musculus	517	52	77.6	383	2	Q60119_SCHPO	Q60119 schizosacch
445	52	77.6	151	1	YWIS_WHEAT	P14723 triticum ae	518	52	77.6	387	2	Q9Z2U3_MOUSE	Q9z2u3 mus musculus
446	52	77.6	151	2	Q41192_NICAL	Q41192 nicotiana a	519	52	77.6	391	2	Q8FXK6_BRUSU	Q8fxk6 bruceella su
447	52	77.6	153	1	PRIMA_HUMAN	Q86xr5 homo sapien	520	52	77.6	393	2	Q9N2Z3_CAEEL	Q9n2z3 caenorhabdi
448	52	77.6	153	1	PRIMA_MOUSE	Q810f0 mus musculus	521	52	77.6	393	2	Q91TI4_TUHV1	Q91ti4 tupaiid her
449	52	77.6	153	1	SORB_PIG	P28220 sus scrofa	522	52	77.6	396	2	Q5QIK4_MAIZE	Q5qlk4 zea mays (m
450	52	77.6	158	2	Q95ME6_PIG	Q95me6 sus scrofa	523	52	77.6	396	2	Q5QIK1_ZEAMP	Q5qlk1 zea mays (s
451	52	77.6	158	2	Q9DVW0_9BACU	Q9dvw0 plutella xy	524	52	77.6	396	2	Q5QIJ3_ZEAMP	Q5qlj3 zea mays (s
452	52	77.6	162	2	Q8XSK0_RALSO	Q8xsk0 ralstonia s	525	52	77.6	399	2	Q86VE0_HUMAN	Q86ve0 homo sapien
453	52	77.6	170	2	Q8XTG7_RALSO	Q8xtg7 ralstonia s	526	52	77.6	402	2	Q9UPV1_HUMAN	Q9upv1 homo sapien
454	52	77.6	176	2	Q16990_ACRDO	Q16990 acropora do	527	52	77.6	408	2	Q710T9_POPDE	Q710t9 populus del
455	52	77.6	179	2	Q9D0Y6_MOUSE	Q9d0y6 mus musculus	528	52	77.6	409	2	Q96RE2_HUMAN	Q96re2 homo sapien
456	52	77.6	182	2	Q5M833_RAT	Q5m833 rattus norv	529	52	77.6	412	2	Q10610_9NUCL	Q10610 helicoverpa
457	52	77.6	186	2	Q00486_9CNID	Q00486 hydra sp. m	530	52	77.6	413	2	Q8V5X7_9NUCL	Q8v5x7 helicoverpa
458	52	77.6	187	2	Q67YU3_ARATH	Q67yu3 arabidopsis	531	52	77.6	413	2	Q99H44_9NUCL	Q99h44 helicoverpa
459	52	77.6	192	2	Q4WN95_ASPFU	Q4wn95 aspergillus	532	52	77.6	418	2	Q5ZIC9_CHICK	Q5zic9 gallus gall
460	52	77.6	197	2	Q55BN3_DICDI	Q55bn3 dictyosteli	533	52	77.6	421	1	ACRO_HUMAN	P10323 homo sapien
461	52	77.6	206	2	Q825Z2_STRAW	Q825z2 streptomyce	534	52	77.6	421	2	Q6ICK2_HUMAN	Q6ick2 homo sapien
462	52	77.6	211	2	Q9M644_ARATH	Q9m644 arabidopsis	535	52	77.6	428	2	Q23370_ARATH	Q23370 arabidopsis
463	52	77.6	217	2	Q8IFN2_PLAF7	Q8ifn2 plasmodium	536	52	77.6	431	1	KNAT3_ARATH	P48000 arabidopsis
464	52	77.6	217	2	Q8MS64_DROME	Q8ms64 drosophila	537	52	77.6	431	2	Q8LDC1_ARATH	Q8ldc1 arabidopsis
465	52	77.6	221	2	Q8N6S7_HUMAN	Q8n6s7 homo sapien	538	52	77.6	442	2	Q8T9S2_DERVA	Q8t9s2 dermacentor
466	52	77.6	232	1	ACROL_HUMAN	P58840 homo sapien	539	52	77.6	450	2	Q7Z3N4_HUMAN	Q7z3n4 homo sapien
467	52	77.6	241	2	Q58WN4_ZEALU	Q58wn4 zea luxuria	540	52	77.6	451	2	Q9SRL3_ARATH	Q9srl3 arabidopsis
468	52	77.6	241	2	Q8R2W2_MOUSE	Q8r2w2 mus musculus	541	52	77.6	452	2	Q4TB69_TETNG	Q4tb69 tetraodon n
469	52	77.6	242	2	Q58WNI_9POAL	Q58wn1 tripsacum a	542	52	77.6	453	2	Q86S65_CAEEL	Q86s65 caenorhabdi

543	52	77.6	455	2	QSAED9_CANAL	Q5aed9 candida alb	616	52	77.6	695	2	Q4R565_MACFA	Q4r565 macaca fasc
544	52	77.6	456	2	Q655Q5_ORYSA	Q655q5 oryza sativ	617	52	77.6	698	2	Q7XF73_ORYSA	Q7xf73 oryza sativ
545	52	77.6	457	2	Q52KK4_RAT	Q52kk4 rattus norv	618	52	77.6	700	2	Q8S976_ORYSA	Q8s976 oryza sativ
546	52	77.6	464	2	Q41645_VOLCA	Q41645 volvox cart	619	52	77.6	700	2	Q653H7_ORYSA	Q653h7 oryza sativ
547	52	77.6	468	2	Q86TK9_HUMAN	Q86tk9 homo sapien	620	52	77.6	707	2	Q5RHD8_BRARE	Q5rhd8 brachydanio
548	52	77.6	471	2	Q503D7_BRARE	Q503d7 brachydanio	621	52	77.6	708	2	Q9LJG1_ARATH	Q9ljg1 arabidopsis
549	52	77.6	473	2	Q51RH2_MAGGR	Q51rh2 magnaporthe	622	52	77.6	708	2	Q8BR71_MOUSE	Q8br71 mus musculu
550	52	77.6	475	2	Q8VDD8_MOUSE	Q8vdd8 mus musculu	623	52	77.6	711	1	UBP51_HUMAN	Q70ek9 homo sapien
551	52	77.6	475	2	Q6Y636_MOUSE	Q6y636 mus musculu	624	52	77.6	711	2	Q7S0F5_NEUCR	Q7s0f5 neurospora
552	52	77.6	477	2	Q851M1_ORYSA	Q851m1 oryza sativ	625	52	77.6	712	2	Q52KF5_MOUSE	Q52kf5 mus musculu
553	52	77.6	482	2	Q6JV11_XENLA	Q6jv11 xenopus lae	626	52	77.6	724	1	KCNN3_PIG	P58392 sus scrofa
554	52	77.6	487	1	EBN2_EBV	P12978 epstein-bar	627	52	77.6	727	2	Q4S6F4_TETNG	Q4s6f4 tetraodon n
555	52	77.6	487	2	Q77H1_9GAMA	Q77h1 human herpe	628	52	77.6	730	2	O45021_CAEL	O45021 caenorhabdi
556	52	77.6	488	1	BTBD1_MOUSE	P58544 mus musculu	629	52	77.6	732	2	Q60XR9_CAEBR	Q60xr9 caenorhabdi
557	52	77.6	488	2	Q5M8N6_MOUSE	Q5m8n6 mus musculu	630	52	77.6	741	2	Q8T9G4_ROME	Q8t9g4 drosophila
558	52	77.6	488	2	Q5RK12_RAT	Q5rk12 rattus norv	631	52	77.6	753	2	Q6H899_ORYSA	Q6h899 oryza sativ
559	52	77.6	488	2	Q6GQU7_MOUSE	Q6gqu7 mus musculu	632	52	77.6	757	2	Q9RVC6_DEIRA	Q9rvc6 deinococcus
560	52	77.6	490	2	Q69023_9GAMA	Q69023 human herpe	633	52	77.6	759	2	Q523F8_MAGGR	Q523f8 magnaporthe
561	52	77.6	493	2	Q9UVD1_PNECA	Q9uvd1 pneumocysti	634	52	77.6	773	2	Q7XHB6_ORYSA	Q7xhb6 oryza sativ
562	52	77.6	494	2	Q9FXA1_ARATH	Q9fxa1 arabidopsis	635	52	77.6	773	2	Q8S7Y7_ORYSA	Q8s7y7 oryza sativ
563	52	77.6	494	2	Q6ME19_PARUW	Q6me19 parachlamyd	636	52	77.6	774	2	Q868B4_CAEL	Q868b4 caenorhabdi
564	52	77.6	500	2	Q5QLD5_ORYSA	Q5qld5 oryza sativ	637	52	77.6	787	2	O94096_PNECA	O94096 pneumocysti
565	52	77.6	508	2	Q68Y49_ORYSA	Q68y49 oryza sativ	638	52	77.6	795	2	Q6EUQ1_ORYSA	Q6euq1 oryza sativ
566	52	77.6	508	2	Q9SLZ6_ORYSA	Q9slz6 oryza sativ	639	52	77.6	797	2	Q7XHB9_ORYSA	Q7xhb9 oryza sativ
567	52	77.6	518	2	Q8T0I1_DROME	Q8t0i1 drosophila	640	52	77.6	797	2	Q8S7Z0_ORYSA	Q8s7z0 oryza sativ
568	52	77.6	518	2	Q9SEG2_MAIZE	Q9seg2 zea mays (m	641	52	77.6	803	1	PDE4D_RAT	P14270 rattus norv
569	52	77.6	519	2	O48655_ORYSA	O48655 oryza sativ	642	52	77.6	809	1	PDE4D_HUMAN	Q08499 homo sapien
570	52	77.6	520	1	WASP_MOUSE	P70315 mus musculu	643	52	77.6	811	2	Q80T97_MOUSE	Q80t97 mus musculu
571	52	77.6	520	2	Q53WY0_MOUSE	Q53wy0 mus musculu	644	52	77.6	814	2	O4IR23_GIBZE	Q4ir23 gibberella
572	52	77.6	520	2	Q61078_MOUSE	Q61078 mus musculu	645	52	77.6	818	2	Q8C8U8_MOUSE	Q8c8u8 mus musculu
573	52	77.6	526	2	Q5RGW5_BRARE	Q5rgw5 brachydanio	646	52	77.6	820	2	Q69375_9BETA	Q69375 murid herpe
574	52	77.6	540	2	Q7XME8_ORYSA	Q7xme8 oryza sativ	647	52	77.6	826	1	LOZEN_DROME	Q9w349 drosophila
575	52	77.6	542	2	O11448_9NUCL	O11448 anagrapha f	648	52	77.6	836	2	Q8SNA3_MOUSE	Q8sna3 mus musculu
576	52	77.6	542	2	O92377_NPVB	O92377 bombyx mori	649	52	77.6	845	2	Q7S4M4_NEUCR	Q7s4m4 neurospora
577	52	77.6	542	2	Q77WF9_NPVRO	Q77wf9 rachiplusia	650	52	77.6	856	2	Q9LNA8_ARATH	Q9lna8 arabidopsis
578	52	77.6	543	1	VP61_NPVAC	Q03209 autographa	651	52	77.6	863	1	HCN2_MOUSE	O88703 mus musculu
579	52	77.6	551	2	Q52B68_ORYSA	Q52b68 oryza sativ	652	52	77.6	863	2	Q6BCT5_RAT	Q6bct5 rattus norv
580	52	77.6	560	2	Q9H985_HUMAN	Q9h985 homo sapien	653	52	77.6	864	2	Q8H837_ORYSA	Q8h837 oryza sativ
581	52	77.6	570	2	Q8L3T8_ORYSA	Q8l3t8 oryza sativ	654	52	77.6	870	2	Q8W028_ARATH	Q8w028 arabidopsis
582	52	77.6	576	2	Q742S0_MYCPA	Q742s0 mycobacteri	655	52	77.6	874	2	Q6AHU5_PNECA	Q6ahu5 pneumocysti
583	52	77.6	576	2	Q5ZK52_CHICK	Q5zk52 gallus gall	656	52	77.6	874	2	Q96VJ2_PNECA	Q96vj2 pneumocysti
584	52	77.6	577	2	Q7XRW4_ORYSA	Q7xrw4 oryza sativ	657	52	77.6	893	2	Q6ZCX3_ORYSA	Q6zcx3 oryza sativ
585	52	77.6	585	2	Q7XQM8_ORYSA	Q7xqm8 oryza sativ	658	52	77.6	902	2	Q6AHS3_PNECA	Q6ahs3 pneumocysti
586	52	77.6	588	2	Q9SNE9_ARATH	Q9sne9 arabidopsis	659	52	77.6	902	2	Q6AHV9_PNECA	Q6ahv9 pneumocysti
587	52	77.6	593	2	Q8LMC6_ORYSA	Q8lmc6 oryza sativ	660	52	77.6	907	2	Q8LIK9_ORYSA	Q8lik9 oryza sativ
588	52	77.6	600	2	Q8N4J6_HUMAN	Q8n4j6 homo sapien	661	52	77.6	909	2	Q7XMC9_ORYSA	Q7xmc9 oryza sativ
589	52	77.6	605	1	NCKX4_HUMAN	Q8nff2 homo sapien	662	52	77.6	909	2	Q56B20_RICBE	Q56b20 rickettsia
590	52	77.6	605	1	NCKX4_MOUSE	Q8c9q8 mus musculu	663	52	77.6	914	2	Q86YY0_HUMAN	Q5rim8 mus musculu
591	52	77.6	613	1	Q6ZLD1_ORYSA	Q6zld1 oryza sativ	664	52	77.6	914	2	Q86YY0_HUMAN	Q86yy0 homo sapien
592	52	77.6	618	2	Q8S6K9_ORYSA	Q8s6k9 oryza sativ	665	52	77.6	927	2	Q7XSN2_ORYSA	Q7xsn2 oryza sativ
593	52	77.6	637	2	Q4P3H3_USTWA	Q4p3h3 ustilago ma	666	52	77.6	938	2	Q96VI4_PNECA	Q96vi4 pneumocysti
594	52	77.6	639	2	Q6DCV1_XENLA	Q6dcv1 xenopus lae	667	52	77.6	941	2	Q9C6N0_ARATH	Q9c6n0 arabidopsis
595	52	77.6	640	2	Q60593_HUMAN	Q60593 homo sapien	668	52	77.6	942	2	O5YRU1_NOCFA	O5yru1 nocardia fa
596	52	77.6	643	2	Q4WCV2_ASPFU	Q4wcv2 aspergillus	669	52	77.6	947	2	Q6AHT0_PNECA	Q6aht0 pneumocysti
597	52	77.6	645	2	Q96EX0_HUMAN	Q96ex0 homo sapien	670	52	77.6	967	2	Q8TDL2_HUMAN	Q8tdl2 homo sapien
598	52	77.6	645	2	O23374_ARATH	Q23374 arabidopsis	671	52	77.6	984	2	Q6L4M1_ORYSA	Q6l4m1 oryza sativ
599	52	77.6	646	2	Q6PFT9_BRARE	Q6pft9 brachydanio	672	52	77.6	996	2	Q924M9_RAT	Q924m9 rattus norv
600	52	77.6	647	2	Q4U2V7_CHLRE	Q4u2v7 chlamydomon	673	52	77.6	996	2	Q6JTV1_MOUSE	Q6jtv1 mus musculu
601	52	77.6	648	1	DREB_HUMAN	Q16643 homo sapien	674	52	77.6	996	2	Q6P730_RAT	Q6p730 rattus norv
602	52	77.6	651	2	Q9UFZ5_HUMAN	Q9ufz5 homo sapien	675	52	77.6	1010	2	Q8BUZ9_MOUSE	Q8buz9 mus musculu
603	52	77.6	653	2	Q6CUH7_KLULA	Q6cu7 kluyveromyc	676	52	77.6	1012	2	Q58HI1_XENLA	Q58hi1 xenopus lae
604	52	77.6	653	2	Q73TB8_MYCPA	Q73tb8 mycobacteri	677	52	77.6	1013	2	Q9NT81_HUMAN	Q9nt81 homo sapien
605	52	77.6	655	2	Q6P2L6_MOUSE	Q6p2l6 mus musculu	678	52	77.6	1015	2	Q8N3X1_HUMAN	Q8n3x1 homo sapien
606	52	77.6	666	2	Q8I2Z7_HUMAN	Q8izz7 homo sapien	679	52	77.6	1021	2	Q9XUV2_CAEL	Q9xuv2 caenorhabdi
607	52	77.6	666	2	O60592_HUMAN	O60592 homo sapien	680	52	77.6	1026	2	Q5VR02_ORYSA	Q5vr02 oryza sativ
608	52	77.6	666	2	Q7Z5R6_HUMAN	Q7z5r6 homo sapien	681	52	77.6	1029	2	O5SQ40_HUMAN	Q5sq40 homo sapien
609	52	77.6	666	2	Q7TVG2_MYCBO	Q7tv2 mycobacteri	682	52	77.6	1031	2	Q8BNC8_MOUSE	Q8bnc8 mus musculu
610	52	77.6	666	2	O69740_MYCTU	Q69740 mycobacteri	683	52	77.6	1036	2	Q9C0C0_HUMAN	Q9c0c0 homo sapien
611	52	77.6	668	2	Q6Y8G4_RABIT	Q6y8g4 oryctolagus	684	52	77.6	1050	2	Q9Y2L7_HUMAN	Q9y2l7 homo sapien
612	52	77.6	670	2	Q8RVE5_ORYSA	Q8rve5 oryza sativ	685	52	77.6	1052	2	Q6ZQ03_MOUSE	Q6zq03 mus musculu
613	52	77.6	672	2	Q8LMH2_ORYSA	Q8lmh2 oryza sativ	686	52	77.6	1065	2	Q96SE1_HUMAN	Q96se1 homo sapien
614	52	77.6	684	2	Q9S282_STRCO	Q9s282 streptomyce	687	52	77.6	1077	2	Q5SQ42_HUMAN	Q5sq42 homo sapien
615	52	77.6	687	2	Q948Y7_VOLCA	Q948y7 volvox cart	688	52	77.6	1077	2	Q9JHC1_MOUSE	Q9jhc1 mus musculu

616	52	77.6	695	2	Q4R565_MACFA	Q4r565 macaca fasc
617	52	77.6	698	2	Q7XF73_ORYSA	Q7xf73 oryza sativ
618	52	77.6	700	2	Q8S976_ORYSA	Q8s976 oryza sativ
619	52	77.6	700	2	Q653H7_ORYSA	Q653h7 oryza sativ
620	52	77.6	707	2	Q5RHD8_BRARE	Q5rhd8 brachydanio
621	52	77.6	708	2	Q9LJG1_ARATH	Q9ljg1 arabidopsis
622	52	77.6	708	2	Q8BR71_MOUSE	Q8br71 mus musculu
623	52	77.6	711	1	UBP51_HUMAN	Q70ek9 homo sapien
624	52	77.6	711	2	Q7S0F5_NEUCR	Q7s0f5 neurospora
625	52	77.6	712	2	Q52KF5_MOUSE	Q52kf5 mus musculu
626	52	77.6	724	1	KCNN3_PIG	P58392 sus scrofa
627	52	77.6	727	2	Q4S6F4_TETNG	Q4s6f4 tetraodon n
628	52	77.6	730	2	O45021_CAEL	O45021 caenorhabdi
629	52	77.6	732	2	Q60XR9_CAEBR	Q60xr9 caenorhabdi
630	52	77.6	741	2	Q8T9G4_DROME	Q8t9g4 drosophila
631	52	77.6	753	2	Q6H899_ORYSA	Q6h899 oryza sativ
632	52	77.6	757	2	Q9RVC6_DEIRA	Q9rvc6 deinococcus
633	52	77.6	759	2	Q523F8_MAGGR	Q523f8 magnaporthe
634	52	77.6	773	2	Q7XHB6_ORYSA	Q7xhb6 oryza sativ
635	52	77.6	773	2	Q8S7Y7_ORYSA	Q8s7y7 oryza sativ
636	52	77.6	774	2	Q868B4_CAEL	Q868b4 caenorhabdi
637	52	77.6	787	2	O94096_PNECA	O94096 pneumocysti
638	52	77.6	795	2	Q6EUQ1_ORYSA	Q6euq1 oryza sativ
639	52	77.6	797	2	Q7XHB9_ORYSA	Q7xhb9 oryza sativ
640	52	77.6	797	2	Q8S7Z0_ORYSA	Q8s7z0 oryza sativ
641	52	77.6	803	1	PDE4D_RAT	P14270 rattus norv
642	52	77.6	809	1	PDE4D_HUMAN	Q08499 homo sapien
643	52	77.6	811	2	Q80T97_MOUSE	Q80t97 mus musculu
644	52	77.6	814	2	Q4IR23_GIBZE	Q4ir23 gibberella
645	52	77.6	818	2	Q8C8U8_MOUSE	Q8c8u8 mus musculu
646	52	77.6	820	2	Q69375_9BETA	Q69375 murid herpe
647	52	77.6	826	1	LOZEN_DROME	Q9w349 drosophila
648	52	77.6	836	2	Q8SNA3_MOUSE	Q8sna3 mus musculu
649	52	77.6	845	2	Q7S4M4_NEUCR	Q7s4m4 neurospora
650	52	77.6	856	2	Q9LNA8_ARATH	Q9lna8 arabidopsis
651	52	77.6	863	1	HCN2_MOUSE	Q88703 mus musculu
652	52	77.6	863	2	Q6BCT5_RAT	Q6bct5 rattus norv
653	52	77.6	864	2	Q8H837_ORYSA	Q8h837 oryza sativ
654	52	77.6	870	2	Q8W028_ARATH	Q8w028 arabidopsis
655	52	77.6	874	2	Q6AHU5_PNECA	Q6ahu5 pneumocysti
656	52	77.6	874	2	Q96VJ2_PNECA	Q96vj2 pneumocysti
657	52	77.6	893	2	Q6ZCX3_ORYSA	Q6zcx3 oryza sativ
658	52	77.6	902	2	Q6AHS3_PNECA	Q6ahs3 pneumocysti
659	52	77.6	902	2	Q6AHV9_PNECA	Q6ahv9 pneumocysti
660	52	77.6	907	2	Q8LIK9_ORYSA	Q8lik9 oryza sativ
661	52	77.6	909	2	Q7XMC9_ORYSA	Q7xmc9 oryza sativ
662	52	77.6	909	2	Q56B20_RICBE	Q56b20 rickettsia
663	52	77.6	910	2	Q5RIM8_MOUSE	Q5rim8 mus musculu
664	52	77.6	914	2	Q86XY0_HUMAN	Q86yy0 homo sapien
665	52	77.6	927	2	Q7XSN2_ORYSA	Q7xsn2 oryza sativ
666	52	77.6	938	2	Q96VI4_PNECA	Q96vi4 pneumocysti
667	52	77.6	941	2	Q9C6N0_ARATH	Q9c6n0 arabidopsis
668	52	77.6	942	2	Q5YRU1_NOCFA	Q5yrul nocardia fa
669	52	77.6	947	2	Q6AHT0_PNECA	Q6aht0 pneumocysti
670	52	77.6	967	2	Q8TDL2_HUMAN	Q8tdl2 homo sapien
671	52	77.6	984	2	Q6L4M1_ORYSA	Q6l4m1 oryza sativ
672	52	77.6	996	2	Q924M9_RAT	Q924m9 rattus norv
673	52	77.6	996	2	Q6JTV1_MOUSE	Q6jtv1 mus musculu
674	52	77.6	996	2	Q6P730_RAT	Q6p730 rattus norv
675	52	77.6	1010	2	Q8BUZ9_MOUSE	Q8buz9 mus musculu
676	52	77.6	1012	2	Q58HI1_XENLA	Q58hi1 xenopus lae
677	52	77.6	1013	2	Q9NT81_HUMAN	Q9nt81 homo sapien
678	52	77.6	1015	2	Q8N3X1_HUMAN	Q8n3x1 homo sapien
679	52	77.6	1021	2	Q9XUV2_CAEL	Q9xuv2 caenorhabdi
680	52	77.6	1026	2	Q5VR02_ORYSA	Q5vr02 oryza sativ
681	52	77.6	1029	2	Q5SQ40_HUMAN	Q5sq40 homo sapien
682	52	77.6	1031	2	Q8BNC8_MOUSE	Q8bnc8 mus musculu
683	52	77.6	1036	2	Q9C0C0_HUMAN	Q9c0c0 homo sapien
684	52	77.6	1050	2	Q9Y2L7_HUMAN	Q9y2l7 homo sapien
685	52	77.6	1052	2	Q6ZQ03_MOUSE	Q6zq03 mus musculu
686	52	77.6	1065	2	Q96SE1_HUMAN	Q96se1 homo sapien
687	52	77.6	1077	2	Q5SQ42_HUMAN	Q5sq42 homo sapien
688	52	77.6	1077	2	Q9JHC1_MOUSE	Q9jhc1 mus musculu

689	52	77.6	1078	2	Q5SQ43_HUMAN	Q5sq43 homo sapien
690	52	77.6	1092	2	Q55JK0_CRYNE	Q55jk0 cryptococcu
691	52	77.6	1096	2	Q6MG49_RAT	Q6mg49 rattus norv
692	52	77.6	1098	2	Q9WTN8_RAT	Q9wt8 rattus norv
693	52	77.6	1099	2	Q5KA01_CRYNE	Q5ka01 cryptococcu
694	52	77.6	1126	2	Q5HYL9_HUMAN	Q5hyl9 homo sapien
695	52	77.6	1126	2	Q96SA6_HUMAN	Q9esa6 homo sapien
696	52	77.6	1126	2	Q9BCN4_HUMAN	Q9bcn4 homo sapien
697	52	77.6	1126	2	Q8IU41_DICDI	Q8iu41 dictyosteli
698	52	77.6	1126	2	Q54SP2_DICDI	Q54sp2 dictyosteli
699	52	77.6	1132	1	BAT3_HUMAN	P46379 homo sapien
700	52	77.6	1132	2	Q5STC1_HUMAN	Q5stc1 homo sapien
701	52	77.6	1133	2	Q872C5_NEUCR	Q872c5 neurospora
702	52	77.6	1144	2	Q7S3B2_NEUCR	Q7s3b2 neurospora
703	52	77.6	1154	2	Q9Z1R2_MOUSE	Q9z1r2 mus musculu
704	52	77.6	1159	2	Q69ZC2_MOUSE	Q69zc2 mus musculu
705	52	77.6	1160	2	Q5VMQ8_HUMAN	Q5vmq8 homo sapien
706	52	77.6	1162	2	Q5SQ39_HUMAN	Q5sq39 homo sapien
707	52	77.6	1170	2	Q4IC26_GIBZE	Q4ic26 gibberella
708	52	77.6	1171	2	Q94875_HUMAN	Q94875 homo sapien
709	52	77.6	1175	1	HCN4_RABIT	Q94875 homo sapien
710	52	77.6	1175	2	Q4RER3_TETNG	Q9tv66 oryctolagus
711	52	77.6	1190	2	Q54MB0_DICDI	Q4rer3 tetraodon n
712	52	77.6	1197	2	Q5W6I0_ORYSA	Q54mb0 dictyosteli
713	52	77.6	1205	2	Q4WQJ5_AS PFU	Q5w6i0 oryza sativ
714	52	77.6	1214	2	Q80T83_MOUSE	Q4wgj5 aspergillus
715	52	77.6	1220	1	VCIP1_ORYSA	Q850t1 oryza sativ
716	52	77.6	1221	1	VCIP1_RAT	Q8cdg3 mus musculu
717	52	77.6	1222	1	VCIP1_HUMAN	Q8cf97 rattus norv
718	52	77.6	1222	2	Q504T4_HUMAN	Q96jh7 homo sapien
719	52	77.6	1229	2	Q95874_HUMAN	Q504t4 homo sapien
720	52	77.6	1229	2	Q5SQ35_HUMAN	Q95874 homo sapien
721	52	77.6	1245	2	Q80T83_MOUSE	Q5sq35 homo sapien
722	52	77.6	1282	2	Q4Q6C2_LEIMA	Q80t83 mus musculu
723	52	77.6	1307	2	Q9LVN1_ARATH	Q4q6c2 leishmania
724	52	77.6	1312	2	Q9NR59_HUMAN	Q9lvn1 arabidopsis
725	52	77.6	1312	2	Q7Z5V7_HUMAN	Q9nr59 homo sapien
726	52	77.6	1325	2	Q96PW1_HUMAN	Q7z5v7 homo sapien
727	52	77.6	1343	2	Q9H7N4_HUMAN	Q96pw1 homo sapien
728	52	77.6	1362	1	BRD4_HUMAN	Q9h7n4 homo sapien
729	52	77.6	1384	2	Q4S827_TETNG	Q60885 homo sapien
730	52	77.6	1386	1	YLPML_MOUSE	Q4s827 tetraodon n
731	52	77.6	1400	1	BRD4_MOUSE	Q9r0i7 mus musculu
732	52	77.6	1449	2	Q4RY48_TETNG	Q9esu6 mus musculu
733	52	77.6	1456	2	Q4NMZ8_9DELT	Q4ry48 tetraodon n
734	52	77.6	1494	2	Q67FY1_HUMAN	Q4nmz8 anaeromyxob
735	52	77.6	1499	2	Q86UU0_HUMAN	Q67fy1 homo sapien
736	52	77.6	1542	1	SETBP_HUMAN	Q86uu0 homo sapien
737	52	77.6	1556	2	Q4Q626_LEIMA	Q9y6x0 homo sapien
738	52	77.6	1569	2	Q5CLH8_CRYHO	Q4q626 leishmania
739	52	77.6	1594	2	Q6PAL7_MOUSE	Q5clh8 cryptospori
740	52	77.6	1595	2	Q5U5N7_MOUSE	Q6pal7 mus musculu
741	52	77.6	1598	2	Q4REK0_TETNG	Q5u5n7 mus musculu
742	52	77.6	1645	2	Q7TP73_RAT	Q4rek0 tetraodon n
743	52	77.6	1649	2	Q8X037_NEUCR	Q7tp73 rattus norv
744	52	77.6	1662	2	Q7RWF2_NEUCR	Q8x037 neurospora
745	52	77.6	1707	1	SET1_HUMAN	Q7rwf2 neurospora
746	52	77.6	1778	2	Q8IWR5_HUMAN	Q15047 homo sapien
747	52	77.6	1779	2	Q53LC9_ORYSA	Q8iwr5 homo sapien
748	52	77.6	1799	2	Q7XU11_ORYSA	Q53lc9 oryza sativ
749	52	77.6	1819	2	Q4TT23_9SPHN	Q7xul1 oryza sativ
750	52	77.6	1858	2	Q8I2D2_HUMAN	Q4tt23 erythrobact
751	52	77.6	1858	2	Q8NFF8_HUMAN	Q8izd2 homo sapien
752	52	77.6	1859	2	Q4P6X2_USTMA	Q8nff8 homo sapien
753	52	77.6	1893	2	Q7ZWN3_XENLA	Q4p6x2 ustilago ma
754	52	77.6	1938	2	Q66J90_XENLA	Q7zwn3 xenopus lae
755	52	77.6	2008	2	Q5FP38_CHICK	Q66j90 xenopus lae
756	52	77.6	2055	2	Q5NCG1_MOUSE	Q5f3p8 gallus gall
757	52	77.6	2123	2	Q9LPK3_ARATH	Q5ncg1 mus musculu
758	52	77.6	2195	2	Q4PFH2_USTMA	Q9lpk3 arabidopsis
759	52	77.6	2252	2	Q9VZX4_DROME	Q4pfh2 ustilago ma
760	52	77.6	2322	2	Q6UDW6_PLAFA	Q9vzx4 drosophila
761	52	77.6	2330	2	Q7SDG4_NEUCR	Q6udw6 plasmodium

762	52	77.6	2335	2	Q5RIM5_MOUSE	Q5rim5 mus musculu
763	52	77.6	2341	2	Q5RIM7_MOUSE	Q5rim7 mus musculu
764	52	77.6	2343	2	Q7Z516_HUMAN	Q7z516 homo sapien
765	52	77.6	2387	2	Q5PRE9_HUMAN	Q5pre9 mus musculu
766	52	77.6	2398	2	Q4Q3E5_LEIMA	Q4q3e5 leishmania
767	52	77.6	2440	1	NCOR1_HUMAN	Q75376 homo sapien
768	52	77.6	2453	1	NCOR1_MOUSE	Q60974 mus musculu
769	52	77.6	2454	2	Q5RIM6_MOUSE	Q5rim6 mus musculu
770	52	77.6	2601	2	Q4QGR7_LEIMA	Q4qgr7 leishmania
771	52	77.6	2645	2	Q6XK19_DROME	Q6xk19 drosophila
772	52	77.6	2646	2	Q5CLM1_CRYHO	Q5clm1 cryptospori
773	52	77.6	2706	2	O15870_PLAFA	O15870 plasmodium
774	52	77.6	3004	2	Q4SN41_TETNG	O15870 tetraodon n
775	52	77.6	3110	1	HD_RAT	P51111 rattus norv
776	52	77.6	3119	2	Q8IHM0_PLAF7	Q8ihm0 plasmodium
777	52	77.6	3139	2	Q9GM99_PIG	Q9gm99 sus scrofa
778	52	77.6	3795	2	Q4Q754_LEIMA	Q4q754 leishmania
779	51	76.1	57	2	Q13697_HUMAN	Q13697 homo sapien
780	51	76.1	70	2	Q9MAV4_ARATH	Q9mav4 arabidopsis
781	51	76.1	88	2	Q4YAE9_PLABE	Q4yae9 plasmodium
782	51	76.1	107	2	Q6YYSI_ORYSA	Q6yys1 oryza sativ
783	51	76.1	108	2	Q91AX2_9PAPI	Q91ax2 human papil
784	51	76.1	127	2	Q8VAY0_WSSV	Q8vay0 white spot
785	51	76.1	134	2	Q9XIP3_ARATH	Q9xip3 arabidopsis
786	51	76.1	137	2	Q69U86_ORYSA	Q69u86 oryza sativ
787	51	76.1	138	2	Q84TS2_ORYSA	Q84ts2 oryza sativ
788	51	76.1	151	2	Q7UE75_RHOBA	Q7ue75 rhodopirell
789	51	76.1	151	2	Q7UEG3_RHOBA	Q7ueg3 rhodopirell
790	51	76.1	153	2	Q67UD0_ORYSA	Q67ud0 oryza sativ
791	51	76.1	157	2	Q7XV56_ORYSA	Q7xv56 oryza sativ
792	51	76.1	175	2	Q6Z4V0_ORYSA	Q6z4v0 oryza sativ
793	51	76.1	182	2	Q86F22_SCHJA	Q86f22 schistosoma
794	51	76.1	184	2	Q7XX24_ORYSA	Q7xx24 oryza sativ
795	51	76.1	185	2	Q9RZ72_DEIRA	Q9rz72 deinococcus
796	51	76.1	200	2	Q6YTD9_ORYSA	Q6ytd9 oryza sativ
797	51	76.1	200	2	Q9CW04_MOUSE	Q9cw04 mus musculu
798	51	76.1	205	2	Q5CKE2_CRYHO	Q5cke2 cryptospori
799	51	76.1	205	2	Q4NRN0_9DELT	Q4nrn0 anaeromyxob
800	51	76.1	208	2	Q61TA7_CAEBR	Q61ta7 caenorhabdi
801	51	76.1	210	2	Q20327_CAEBL	Q20327 caenorhabdi
802	51	76.1	215	2	Q5Z9C7_ORYSA	Q5z9c7 oryza sativ
803	51	76.1	221	2	O49946_SOLTU	O49946 solanum tub
804	51	76.1	223	2	Q7XJ22_ORYSA	Q7xj22 oryza sativ
805	51	76.1	234	2	Q5PXG5_EUPES	Q5pxg5 euphorbia e
806	51	76.1	249	2	Q6PI38_HUMAN	Q6pi38 homo sapien
807	51	76.1	263	2	Q8ZYDI_PYRAB	Q8zyd1 pyrobaculum
808	51	76.1	273	2	Q88AW2_PSESM	Q88aw2 pseudomonas
809	51	76.1	274	2	Q500G4_PSESY	Q500g4 pseudomonas
810	51	76.1	277	2	Q9A5J8_CAUCR	Q9a5j8 caulobacter
811	51	76.1	278	1	TNFL6_RAT	P36940 rattus norv
812	51	76.1	281	1	ATF5_RAT	Q6p788 rattus norv
813	51	76.1	283	1	ATF5_MOUSE	O70191 mus musculu
814	51	76.1	287	2	Q69X10_ORYSA	Q69x10 oryza sativ
815	51	76.1	289	1	HXD8_MOUSE	P23463 mus musculu
816	51	76.1	290	2	Q9VC76_DROME	Q9vc76 drosophila
817	51	76.1	294	2	Q5VQC4_ORYSA	Q5vqc4 oryza sativ
818	51	76.1	306	2	P93845_PEA	P93845 pisum sativ
819	51	76.1	311	2	Q8LAX9_ARATH	Q8lax9 arabidopsis
820	51	76.1	311	2	Q9M9F0_ARATH	Q9m9f0 arabidopsis
821	51	76.1	312	2	Q8BZD7_MOUSE	Q8bzd7 mus musculu
822	51	76.1	315	1	HA22H_ARATH	Q8lem6 arabidopsis
823	51	76.1	318	2	Q23352_CAEBL	Q23352 caenorhabdi
824	51	76.1	319	2	Q6BQZ9_DEBHA	Q6bqz9 debaryomyce
825	51	76.1	327	2	Q8BHG7_MOUSE	Q8bhg7 m mus muscu
826	51	76.1	328	2	Q5CHL3_CRYHO	Q5chl3 cryptospori
827	51	76.1	338	2	Q6P0M8_HUMAN	Q6p0m8 homo sapien
828	51	76.1	345	2	Q585B2_9TRYP	Q585b2 trypanosoma
829	51	76.1	347	2	Q9Y6U7_HUMAN	Q9y6u7 homo sapien
830	51	76.1	348	2	Q8YJL9_BRUMB	Q8yj19 brucella me
831	51	76.1	354	1	NKX24_MOUSE	Q9eqm3 mus musculu
832	51	76.1	362	1	GAW2_HUMAN	Q9nrz4 homo sapien
833	51	76.1	374	2	Q66YD1_CHLRE	Q66yd1 chlamydomon
834	51	76.1	378	2	Q61JL5_CAEBR	Q61jl5 caenorhabdi

835	51	76.1	391	2	Q42532_ARATH	Q42532 arabidopsis	908	51	76.1	791	2	Q6ATQ3_ORYSA	Q6atq3 oryza sativ
836	51	76.1	391	2	Q60AS3_METCA	Q60as3 methylococc	909	51	76.1	792	2	Q54BJ4_DICDI	Q54bj4 dictyosteli
837	51	76.1	393	2	Q40552_TOBAC	Q40552 nicotiana t	910	51	76.1	796	2	Q8S6M0_ORYSA	Q8s6m0 oryza sativ
838	51	76.1	403	2	Q8BIH7_MOUSE	Q8bih7 mus musculu	911	51	76.1	807	2	Q5DTU6_MOUSE	Q5dtu6 mus musculu
839	51	76.1	412	2	Q5N8V9_ORYSA	Q5n8v9 oryza sativ	912	51	76.1	814	2	Q8LM34_ORYSA	Q8lm34 oryza sativ
840	51	76.1	413	2	Q9A4G2_CAUCR	Q9a4g2 caulobacter	913	51	76.1	814	2	Q9LY50_ARATH	Q9ly50 arabidopsis
841	51	76.1	415	2	Q873H2_NEUCR	Q873h2 neurospora	914	51	76.1	818	2	Q6K289_ORYSA	Q6k289 oryza sativ
842	51	76.1	415	2	Q29015_9CETA	Q29015 sus sp. pre	915	51	76.1	822	2	Q6L4I3_ORYSA	Q6l4i3 oryza sativ
843	51	76.1	418	2	Q60F43_ORYSA	Q60f43 oryza sativ	916	51	76.1	824	2	Q4RID3_TETNG	Q4rid3 tetraodon n
844	51	76.1	422	2	Q5Z5I3_ORYSA	Q5z5i3 oryza sativ	917	51	76.1	833	2	Q7XUV2_ORYSA	Q7xuv2 oryza sativ
845	51	76.1	429	2	Q4VXV5_HUMAN	Q4vxv5 homo sapien	918	51	76.1	835	2	Q9SDN5_TOBAC	Q9sdn5 nicotiana t
846	51	76.1	432	2	Q4VXV6_HUMAN	Q4vxv6 homo sapien	919	51	76.1	849	2	Q8H8K7_ORYSA	Q8h8k7 oryza sativ
847	51	76.1	433	2	Q4U2V8_CHLRE	Q4u2v8 chlamydomon	920	51	76.1	856	2	Q9VR54_DROME	Q9vr54 drosophila
848	51	76.1	439	2	Q5S5O8_CRYNE	Q5s5o8 cryptococcu	921	51	76.1	864	2	Q6YW44_ORYSA	Q6yw44 oryza sativ
849	51	76.1	439	2	Q5KGG0_CRYNE	Q5kgg0 cryptococcu	922	51	76.1	881	2	Q6H7U3_ORYSA	Q6h7u3 oryza sativ
850	51	76.1	440	2	Q01672_ASCSU	Q01672 ascaris suu	923	51	76.1	918	2	Q4SK19_TETNG	Q4sk19 tetraodon n
851	51	76.1	446	2	Q22458_CHLRE	Q22458 chlamydomon	924	51	76.1	978	2	Q4SXP7_TETNG	Q4sxp7 tetraodon n
852	51	76.1	448	1	WT1_RAT	P49952 rattus norv	925	51	76.1	1001	2	Q8VDG6_MOUSE	Q8vdg6 mus musculu
853	51	76.1	449	1	WT1_HUMAN	P19544 homo sapien	926	51	76.1	1006	2	Q9LMQ1_ARATH	Q9lmq1 arabidopsis
854	51	76.1	449	1	WT1_MOUSE	P22561 mus musculu	927	51	76.1	1027	2	Q9W3U2_DROME	Q9w3u2 drosophila
855	51	76.1	449	1	WT1_PIG	Q62651 sus scrofa	928	51	76.1	1067	2	Q6NSN0_BRARE	Q6nsn0 brachydanio
856	51	76.1	451	2	Q6EQW6_ORYSA	Q6eqw6 oryza sativ	929	51	76.1	1165	2	Q6ZHK2_ORYSA	Q6zhk2 oryza sativ
857	51	76.1	453	2	Q4VA17_MOUSE	Q4va17 mus musculu	930	51	76.1	1248	2	Q6S589_ARATH	Q6s589 arabidopsis
858	51	76.1	455	2	Q93XP5_ORYSA	Q93xp5 oryza sativ	931	51	76.1	1271	2	Q5FWL5_XENLA	Q5fwl5 xenopus lae
859	51	76.1	455	2	Q5Z5I4_ORYSA	Q5z5i4 oryza sativ	932	51	76.1	1293	2	Q6CAT2_YARLI	Q6cat2 yarrowia li
860	51	76.1	460	2	Q5N9Q6_ORYSA	Q5n9q6 oryza sativ	933	51	76.1	1392	2	Q9XER9_ARATH	Q9xer9 arabidopsis
861	51	76.1	484	2	Q5HZ98_XENLA	Q5hz98 xenopus lae	934	51	76.1	1422	2	Q53KD8_ORYSA	Q53kd8 oryza sativ
862	51	76.1	485	1	SSGP_VOLCA	P21997 volvox cart	935	51	76.1	1424	2	Q5XJV6_MOUSE	Q5xjv6 mus musculu
863	51	76.1	496	2	Q7QE74_ANOGA	Q7qe74 anopheles g	936	51	76.1	1424	2	Q52KF1_MOUSE	Q52kf1 mus musculu
864	51	76.1	497	1	WASF2_MOUSE	Q8bh43 mus musculu	937	51	76.1	1456	2	Q7SBI9_NEUCR	Q7sbi9 neurospora
865	51	76.1	497	2	Q9FP66_ORYSA	Q9fp66 oryza sativ	938	51	76.1	1480	2	Q96Q04_HUMAN	Q96q04 homo sapien
866	51	76.1	497	2	Q5FWU0_RAT	Q5fwu0 rattus norv	939	51	76.1	1486	2	Q54HS3_DICDI	Q54hs3 dictyosteli
867	51	76.1	498	1	WASF2_HUMAN	Q9y6w5 homo sapien	940	51	76.1	1509	1	GSCR1_HUMAN	Q9nzm4 homo sapien
868	51	76.1	501	2	Q69N33_ORYSA	Q69n33 oryza sativ	941	51	76.1	1568	2	Q5TKD9_ORYSA	Q5tkd9 oryza sativ
869	51	76.1	502	2	Q8NFC6_HUMAN	Q8nfc6 homo sapien	942	51	76.1	1689	2	Q5TQV0_ANOGA	Q5tqv0 anopheles g
870	51	76.1	503	2	Q66IV2_XENLA	Q66iv2 xenopus lae	943	51	76.1	1698	2	Q7Q5B6_ANOGA	Q7q5b6 anopheles g
871	51	76.1	507	2	Q69XH2_ORYSA	Q69xh2 oryza sativ	944	51	76.1	1734	2	Q8MKW9_DROME	Q8mkw9 drosophila
872	51	76.1	507	2	Q8FY77_BRUSU	Q8fy77 brucella su	945	51	76.1	1743	2	Q966V0_DROME	Q966v0 drosophila
873	51	76.1	509	2	Q9GS27_HIRME	Q9gs27 hirudo medi	946	51	76.1	1814	2	Q4PBT2_USTMA	Q4pbt2 ustilago ma
874	51	76.1	516	2	Q48708_ARATH	Q48708 arabidopsis	947	51	76.1	1884	2	Q5CS67_CRYPV	Q5cs67 cryptospori
875	51	76.1	522	2	Q4RS34_TETNG	Q4rs34 tetraodon n	948	51	76.1	1931	2	Q4Q2X5_LEIMA	Q4q2x5 leishmania
876	51	76.1	524	2	Q9FN55_ARATH	Q9fn55 arabidopsis	949	51	76.1	1933	2	Q7EY18_ORYSA	Q7ey18 oryza sativ
877	51	76.1	528	2	Q89PV5_BRAJA	Q89pv5 bradyrhizob	950	51	76.1	1946	2	Q8N3R5_HUMAN	Q8n3r5 homo sapien
878	51	76.1	530	2	Q07701_CHV12	Q07701 cercopithec	951	51	76.1	1997	2	Q8LRM7_CHLRE	Q8lrm7 chlamydomon
879	51	76.1	536	2	Q51VQ7_MAGGR	Q51vq7 magnaporthec	952	51	76.1	2162	2	Q5SEI2_DICDI	Q5sei2 dictyosteli
880	51	76.1	555	1	GP1_CHLRE	Q9fpq6 chlamydomon	953	51	76.1	2274	2	Q54W13_DICDI	Q54wl3 dictyosteli
881	51	76.1	567	2	Q6IDZ1_ANOGA	Q6idz1 anopheles g	954	51	76.1	2376	2	Q966V1_DROME	Q966v1 drosophila
882	51	76.1	570	2	Q4TPN9_9SPHN	Q4tpn9 erythrobact	955	51	76.1	2376	2	Q9V5J0_DROME	Q9v5j0 drosophila
883	51	76.1	571	2	Q9RXG3_DEIRA	Q9rxg3 deinococcus	956	51	76.1	2386	1	EGFL4_HUMAN	Q7z7m0 homo sapien
884	51	76.1	584	2	Q82DU7_STRAW	Q82du7 streptomyce	957	51	76.1	2443	2	Q5CSU8_CRYPV	Q5csu8 cryptospori
885	51	76.1	599	2	P93797_VOLCA	P93797 volvox cart	958	51	76.1	2443	2	Q5CNM5_CRYHO	Q5cnm5 cryptospori
886	51	76.1	601	2	Q5BU15_FUGRU	Q5bul5 fugu rubrip	959	51	76.1	3251	2	Q8K4E0_MOUSE	Q8k4e0 mus musculu
887	51	76.1	606	2	Q852P0_VOLCA	Q852p0 volvox cart	960	51	76.1	3302	2	Q71LX6_RAT	Q71lx6 rattus norv
888	51	76.1	607	2	Q9LU79_ARATH	Q9lu79 arabidopsis	961	51	76.1	3345	2	Q5MJ67_HUMAN	Q5mj67 homo sapien
889	51	76.1	610	2	Q8GY16_ARATH	Q8gy16 arabidopsis	962	51	76.1	5085	1	PCLO_RAT	Q9jks6 rattus norv
890	51	76.1	610	2	Q9MAX4_ARATH	Q9max4 arabidopsis	963	50.5	75.4	437	2	Q9STN0_ARATH	Q9stn0 arabidopsis
891	51	76.1	610	2	Q9SU56_ARATH	Q9su56 arabidopsis	964	50.5	75.4	639	2	Q55VN4_CRYNE	Q55vn4 cryptococcu
892	51	76.1	612	2	Q4SVK0_TETNG	Q4svk0 tetraodon n	965	50.5	75.4	639	2	Q5KKK5_CRYNE	Q5kkk5 cryptococcu
893	51	76.1	622	2	Q7TFD9_RHCM6	Q7tfd9 rhesus cyto	966	50	74.6	20	2	Q7Z5Z9_HUMAN	Q7z5z9 homo sapien
894	51	76.1	625	1	WASF4_HUMAN	Q8iv90 homo sapien	967	50	74.6	33	2	O02832_9PRIM	O02832 gorilla gor
895	51	76.1	625	2	Q8S9B5_VOLCA	Q8s9b5 volvox cart	968	50	74.6	62	2	Q6Z053_ORYSA	Q6z053 oryza sativ
896	51	76.1	654	2	Q51LT9_MAGGR	Q51lt9 magnaporthec	969	50	74.6	67	2	Q9Z1I8_RAT	Q9z1i8 rattus norv
897	51	76.1	675	2	Q8BI66_MOUSE	Q8bi66 mus musculu	970	50	74.6	67	2	Q75PL3_NUMME	Q75pl3 numida mele
898	51	76.1	684	2	Q7Y1D3_ORYSA	Q7y1d3 oryza sativ	971	50	74.6	76	2	Q8LG30_ARATH	Q8lg30 arabidopsis
899	51	76.1	685	2	Q68A65_RAT	Q68a65 rattus norv	972	50	74.6	85	2	Q6YRL7_ORYSA	Q6yrl7 oryza sativ
900	51	76.1	701	2	Q7RWB1_NEUCR	Q7rwb1 neurospora	973	50	74.6	85	2	Q6ZLK9_ORYSA	Q6zlk9 oryza sativ
901	51	76.1	703	2	Q7S872_NEUCR	Q7s872 neurospora	974	50	74.6	89	2	Q8QNL0_9PHYC	Q8qnl0 ectocarpus
902	51	76.1	712	2	Q5N7W0_ORYSA	Q5n7w0 oryza sativ	975	50	74.6	91	2	Q8ZT81_PYRAE	Q8zt81 pyrobaculum
903	51	76.1	715	1	AT12_HHV1F	P08314 human herpe	976	50	74.6	97	2	Q8MZM1_DROVI	Q8mzm1 drosophila
904	51	76.1	738	2	Q7PYV4_ANOGA	Q7pyv4 anopheles g	977	50	74.6	109	2	Q9BLB3_CAEEL	Q9blb3 caenorhabdi
905	51	76.1	759	2	Q8HXV1_RABIT	Q8hxxv1 cryctolagus	978	50	74.6	117	2	Q5SJJG1_THET8	Q5sjgj thermus the
906	51	76.1	760	2	Q9TOK5_ARATH	Q9tok5 arabidopsis	979	50	74.6	118	2	Q01614_PNECA	Q01614 pneumocysti
907	51	76.1	790	2	Q4KSZ9_9NUCL	Q4ksz9 chrysodeixi	980	50	74.6	120	2	Q5A9M7_CANAL	Q5a9m7 candida alb

Q42532 arabidopsis	908	51	76.1	791	2	Q6ATQ3_ORYSA
Q60as3 methylococc	909	51	76.1	792	2	Q54BJ4_DICDI
Q40552 nicotiana t	910	51	76.1	796	2	Q8S6M0_ORYSA
Q8bih7 mus musculu	911	51	76.1	807	2	Q5DTU6_MOUSE
Q5n8v9 oryza sativ	912	51	76.1	814	2	Q8LM34_ORYSA
Q9a4g2 caulobacter	913	51	76.1	814	2	Q9LY50_ARATH
Q873h2 neurospora	914	51	76.1	818	2	Q6K289_ORYSA
Q29015 sus sp. pre	915	51	76.1	822	2	Q6L4I3_ORYSA
Q60f43 oryza sativ	916	51	76.1	824	2	Q4RID3_TETNG
Q5z5i3 oryza sativ	917	51	76.1	833	2	Q7XUV2_ORYSA
Q4vxv5 homo sapien	918	51	76.1	835	2	Q9SDN5_TOBAC
Q4vxv6 homo sapien	919	51	76.1	849	2	Q8H8K7_ORYSA
Q4u2v8 chlamydomon	920	51	76.1	856	2	Q9VR54_DROME
Q5s5o8 cryptococcu	921	51	76.1	864	2	Q6YW44_ORYSA
Q5kgg0 cryptococcu	922	51	76.1	881	2	Q6H7U3_ORYSA
O01672 ascaris suu	923	51	76.1	918	2	Q4SK19_TETNG
Q22458 chlamydomon	924	51	76.1	978	2	Q4SXP7_TETNG
P49952 rattus norv	925	51	76.1	1001	2	Q8VDG6_MOUSE
P19544 homo sapien	926	51	76.1	1006	2	Q9LMQ1_ARATH
P22561 mus musculu	927	51	76.1	1027	2	Q9W3U2_DROME
Q62651 sus scrofa	928	51	76.1	1067	2	Q6NSN0_BRARE
Q6eqw6 oryza sativ	929	51	76.1	1165	2	Q6ZHK2_ORYSA
Q4va17 mus musculu	930	51	76.1	1248	2	Q6S589_ARATH
Q93xp5 oryza sativ	931	51	76.1	1271	2	Q5FWL5_XENLA
Q5z5i4 oryza sativ	932	51	76.1	1293	2	Q6CAT2_YARLI
Q5n9q6 oryza sativ	933	51	76.1	1392	2	Q9XER9_ARATH
Q5hz98 xenopus lae	934	51	76.1	1422	2	Q53KD8_ORYSA
P21997 volvox cart	935	51	76.1	1424	2	Q5XJV6_MOUSE
Q7qe74 anopheles g	936	51	76.1	1424	2	Q52KP1_MOUSE
Q8bh43 mus musculu	937	51	76.1	1456	2	Q7SBI9_NEUCR
Q9fp66 oryza sativ	938	51	76.1	1480	2	Q96Q04_HUMAN
Q5fwu0 rattus norv	939	51	76.1	1486	2	Q54HS3_DICDI
Q9y6w5 homo sapien	940	51	76.1	1509	1	GSCR1_HUMAN
Q69n33 oryza sativ	941	51	76.1	1568	2	Q5TKD9_ORYSA
Q8nfc6 homo sapien	942	51	76.1	1689	2	Q5TQV0_ANOGA
Q66iv2 xenopus lae	943	51	76.1	1698	2	Q7Q5B6_ANOGA
Q69xh2 oryza sativ	944	51	76.1	1734	2	Q8MKW9_DROME
Q8fy77 brucella su	945	51	76.1	1743	2	Q966V0_DROME
Q9gs27 hirudo medi	946	51	76.1	1814	2	Q4PBT2_USTMA
O48708 arabidopsis	947	51	76.1	1884	2	Q5CS67_CRYPV
Q4rs34 tetraodon n	948	51	76.1	1931	2	Q4Q2X5_LEIMA
Q9fn55 arabidopsis	949	51	76.1	1933	2	Q7EY18_ORYSA
Q89pv5 bradyrhizob	950	51	76.1	1946	2	Q8N3R5_HUMAN
Q07701 cercopithec	951	51	76.1	1997	2	Q8LRM7_CHLRE
Q51vq7 magnaporthe	952	51	76.1	2162	2	Q55E12_DICDI
Q9fpq6 chlamydomon	953	51	76.1	2274	2	Q54WL3_DICDI
Q6idz1 anopheles g	954	51	76.1	2376	2	Q966V1_DROME
Q4tpn9 erythrobact	955	51	76.1	2376	2	Q9V5J0_DROME
Q9rxg3 deinococcus	956	51	76.1	2386	1	EGFL4_HUMAN
Q82du7 streptomyce	957	51	76.1	2443	2	Q5CSU8_CRYPV
P93797 volvox cart	958	51	76.1	2443	2	Q5CNM5_CRYHO
Q5bul5 fugu rubrip	959	51	76.1	3251	2	Q8K4E0_MOUSE
Q852p0 volvox cart	960	51	76.1	3302	2	Q71LX6_RAT
Q9lu79 arabidopsis	961	51	76.1	3345	2	Q5MJ67_HUMAN
Q8gy16 arabidopsis	962	51	76.1	5085	1	PCLO_RAT
Q9max4 arabidopsis	963	50.5	75.4	437	2	Q9STN0_ARATH
Q9su56 arabidopsis	964	50.5	75.4	639	2	Q55VN4_CRYNE
Q4svk0 tetraodon n	965	50.5	75.4	639	2	Q5KKK5_CRYNE
Q7tfd9 rhesus cyto	966	50	74.6	20	2	Q72529_HUMAN
Q8iv90 homo sapien	967	50	74.6	33	2	Q02832_9PRIM
Q8s9b5 volvox cart	968	50	74.6	62	2	Q6Z053_ORYSA
Q51lt9 magnaporthe	969	50	74.6	67	2	Q9Z1I8_RAT
Q8bi66 mus musculu	970	50	74.6	67	2	Q75PL3_NUMME
Q7yld3 oryza sativ	971	50	74.6	76	2	Q8LG30_ARATH
Q68a65 rattus norv	972	50	74.6	85	2	Q6YRL7_ORYSA
Q7rwb1 neurospora	973	50	74.6	85	2	Q6ZLK9_ORYSA
Q7s872 neurospora	974	50	74.6	89	2	Q8QNL0_9PHYC
Q5n7w0 oryza sativ	975	50	74.6	91	2	Q8ZT81_PYRAE
P08314 human herpe	976	50	74.6	97	2	Q8MZM1_DROVI
Q7pyv4 anopheles g	977	50	74.6	109	2	Q9BLB3_CAEEL
Q8hvx1 oryctolagus	978	50	74.6	117	2	Q5SJG1_THET8
Q9t0k5 arabidopsis	979	50	74.6	118	2	Q01614_PNECA
Q4ksz9 chrysodeixi	980	50	74.6	120	2	Q5A9M7_CANAL

981 50 74.6 121 2 Q5A9C9_CANAL Q5a9c9 candida alb
982 50 74.6 124 2 Q9FN57_ARATH Q9fn57 arabidopsis
983 50 74.6 126 2 Q8H604_ORYSA Q8h604 oryza sativ
984 50 74.6 127 2 Q4THL7_TETNG Q4thl7 tetraodon n
985 50 74.6 128 2 Q6QNH4_9ORTH Q6qnh4 chorthippus
986 50 74.6 134 2 Q5DGR5_SCHJA Q5dgr5 schistosoma
987 50 74.6 137 2 Q5SPH8_BRARE Q5sph8 brachydanio
988 50 74.6 138 2 Q4UGR6_THEAN Q4ugr6 theileria a
989 50 74.6 141 1 YPRO_OWEFU P21260 owenia fusi
990 50 74.6 142 2 Q00485_9CNID Q00485 hydra sp. m
991 50 74.6 146 2 Q80VV6_MOUSE Q80vv6 mus musculu
992 50 74.6 148 2 Q78DZ0_RATRT Q78dz0 rattus ratt
993 50 74.6 148 2 Q64371_RAT Q64371 rattus norv
994 50 74.6 150 2 Q5D9Q6_SCHJA Q5d9q6 schistosoma
995 50 74.6 150 2 Q6Z237_ORYSA Q6z237 oryza sativ
996 50 74.6 157 2 Q7PNI8_ANOGA Q7pni8 anopheles g
997 50 74.6 158 2 Q6RW62_RAT Q6rw62 rattus norv
998 50 74.6 161 2 O23119_ARATH O23119 arabidopsis
999 50 74.6 161 2 Q9M0N0_ARATH Q9m0n0 arabidopsis
1000 50 74.6 162 2 Q23137_CAEBL Q23137 caenorhabdi

ALIGNMENTS

RESULT 1
Q740X0_MYCPA
ID Q740X0_MYCPA PRELIMINARY; PRT; 480 AA.
AC Q740X0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=MAP1222;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017231; AAS03539.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR003661; His_kinA_N.
DR InterPro; IPR003660; His_kin_HAMP.
DR InterPro; IPR004358; His_kin_like_C.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF02518; HATPase_c; 1.
DR Pfam; PF00512; H1eKA; 1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR SMART; SM00304; HAMP; 1.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00388; H1eKA; 1.
DR PROSITE; PS50885; HAMP; 1.
DR PROSITE; PS50109; HIS_KIN; 1.
KW Complete proteome.
SQ SEQUENCE 480 AA; 50782 MW; C85E333BEAD3181C CRC64;

Query Match 88.1%; Score 59; DB 2; Length 480;
Best Local Similarity 90.0%; Pred. No. 82;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY . 1 PPPYPPPPPIP 10

Db 121 PPPYPPPPPGP 130
RESULT 2
Q6Z1Z7_ORYSA
ID Q6Z1Z7_ORYSA PRELIMINARY; PRT; 347 AA.
AC Q6Z1Z7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE PWWP domain protein-like.
GN Name=B1147B12.7-2;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Katayose Y.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005406; BAD03547.1; -; Genomic_DNA.
DR Gramene; Q6Z1Z7; -.
SQ SEQUENCE 347 AA; 36953 MW; 3250DA6E418E70E8 CRC64;

Query Match 86.6%; Score 58; DB 2; Length 347;
Best Local Similarity 80.0%; Pred. No. 75;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
Db 151 PPPIPPPPPVP 160

RESULT 3
Q6ZTK4_HUMAN
ID Q6ZTK4_HUMAN PRELIMINARY; PRT; 467 AA.
AC Q6ZTK4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein FLJ44573.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Uterus;
RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K.,
RA Isogai T.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK126537; BAC86582.1; -; mRNA.
SQ SEQUENCE 467 AA; 50372 MW; B27AB1A7A25421F0 CRC64;

Query Match 86.6%; Score 58; DB 2; Length 467;
Best Local Similarity 80.0%; Pred. No. 1e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
Db 176 PPAYPPPPPPVP 185

RESULT 4


```
Q6ZVU3 HUMAN
ID Q6ZVU3_HUMAN PRELIMINARY; PRT; 536 AA.
AC Q6ZVU3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein FLJ42079.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Synovial membrane tissue;
RA Suzuki O., Sasaki N., Aotsuka S., Shoji T., Ichihara T., Shiohata N.,
RA Matsumoto K., Hirano M., Sano S., Nomura R., Yoshikawa Y.,
RA Matsumura Y., Moriya S., Chiba E., Momiyama H., Onogawa S.,
RA Kaeriyama S., Satoh N., Matsunawa H., Takahashi E., Kataoka R.,
RA Kuga N., Kuroda A., Satoh I., Kamata K., Takami S., Terashima Y.,
RA Watanabe M., Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A.,
RA Ishii S., Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Yamashita H., Matsuo K., Nakamura Y., Sekine M.,
RA Kikuchi H., Kanda K., Wagatsuma M., Murakawa K., Kanehori K.,
RA Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y.,
RA Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.,
RA Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; AK124073; BAC85767.1; -; mRNA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR000980; SH2.
DR Pfam; PF00017; SH2; 1.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS50001; SH2; 1.
SQ SEQUENCE 536 AA; 58716 MW; F6AF209F37091694 CRC64;

Query Match 86.6%; Score 58; DB 2; Length 536;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
|| |||||:|
DB 176 PPAYPPPPVP 185

RESULT 5
Q5ZHV9 CHICK
ID Q5ZHV9_CHICK PRELIMINARY; PRT; 539 AA.
AC Q5ZHV9;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN ORFNames=RCJMB04_3217;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CB; TISSUE=Bursa;
RA Caldwell R.B., Kierzek A.M., Arakawa H., Bezzubov Y., Zaim J.,
RA Fiedler P., Kutter S., Blagodatski A., Kostovska D., Koter M.,
RA Plachy J., Carninci P., Hayashizaki Y., Buerstedde J.M.;
RT "Full-length cDNAs from chicken bursal lymphocytes to facilitate
RT genefunction analysis.";
RL Genome Biol. 6:R6-R6(2005).
CC -1- SIMILARITY: Contains 1 PH domain.
DR EMBL; AJ721025; CAG32684.1; -; mRNA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000980; SH2.
```

```
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00017; SH2; 1.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS50003; PH DOMAIN; 1.
DR PROSITE; PS50001; SH2; 1.
KW Hypothetical protein; SH2 domain.
SQ SEQUENCE 539 AA; 60676 MW; 028D9154C98E12B6 CRC64;

Query Match 86.6%; Score 58; DB 2; Length 539;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
|| |||||:|
DB 200 PPAYPPPPVP 209

RESULT 6
3BP2 MOUSE
ID 3BP2_MOUSE STANDARD; PRT; 559 AA.
AC Q06649;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE SH3 domain-binding protein 2 (3BP-2).
GN Name=Sh3bp2; Synonyms=3bp2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93174278; PubMed=8438166;
RA Ren R., Mayer B.J., Cicchetti P., Baltimore D.;
RT "Identification of a ten-amino acid proline-rich SH3 binding site.";
RL Science 259:1157-1161(1993).
CC -1- FUNCTION: Binds differentially to the SH3 domains of certain
CC proteins of signal transduction pathways. Binds to
CC phosphatidylinositols; linking the hemopoietic tyrosine kinase fes
CC to the cytoplasmic membrane in a phosphorylation dependent
CC mechanism (By similarity).
CC -1- SIMILARITY: Contains 1 PH domain.
CC -1- SIMILARITY: Contains 1 SH2 domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; L14543; AAA37121.1; -; mRNA.
DR PIR; I49444; I49444.
DR HSSP; Q9HB21; IEAZ.
DR Ensembl; ENSMUSG00000054520; Mus musculus.
DR MGI; MGI:1346349; Sh3bp2.
DR InterPro; IPR001849; PH.
DR InterPro; IPR011993; PH type.
DR InterPro; IPR000980; SH2.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00017; SH2; 1.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS50003; PH DOMAIN; 1.
DR PROSITE; PS50001; SH2; 1.
KW SH2 domain; SH3-binding.
FT DOMAIN 26 130 PH.
FT DOMAIN 455 553 SH2.
FT MOTIF 201 210 SH3-binding.
FT COMPIAS 205 212 Poly-Pro.
```

FT COMPBIAS 236 240 Poly-Pro.
SQ SEQUENCE 559 AA; 62208 MW; EDFEF11B259646E CRC64;

Query Match 86.6%; Score 58; DB 1; Length 559;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPIP 10
Db 201 PPAYPPPPVP 210

RESULT 7
Q5U3L0 MOUSE
ID Q5U3L0_MOUSE PRELIMINARY; PRT; 559 AA.
AC Q5U3L0;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Sh3bp2 protein.
GN Name=Sh3bp2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Director MGC Project;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 PH domain.
DR EMBL; BC085497; AAH85497.1; -; mRNA.
DR MGI; MGI:1346349; Sh3bp2.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000980; SH2.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00017; SH2; 1.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS50003; PH DOMAIN; 1.
DR PROSITE; PS50001; SH2; 1.
KW SH2 domain.
SQ SEQUENCE 559 AA; 62188 MW; 62235DDE3A849E79 CRC64;

Query Match 86.6%; Score 58; DB 2; Length 559;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;

Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPIP 10
Db 201 PPAYPPPPVP 210

RESULT 8
Q91Z52 MOUSE
ID Q91Z52_MOUSE PRELIMINARY; PRT; 559 AA.
AC Q91Z52;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE SH3-domain binding protein 2.
GN Name=Sh3bp2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FVB/N;
RX TISSUE=Mammary tumor. Metallothionien-TGF alpha model. 10 month old
virgin mouse. Taken by biopsy.;
MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FVB/N;
RC TISSUE=Mammary tumor. Metallothionien-TGF alpha model. 10 month old
virgin mouse. Taken by biopsy.;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 PH domain.
DR EMBL; BC010198; AAH10198.1; -; mRNA.
DR Ensembl; ENSMUSG00000054520; Mus musculus.
DR MGI; MGI:1346349; Sh3bp2.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR001849; PH.
DR InterPro; IPR011993; PH type.
DR InterPro; IPR000980; SH2.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00017; SH2; 1.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS50003; PH DOMAIN; 1.
DR PROSITE; PS50001; SH2; 1.
KW SH2 domain.
SQ SEQUENCE 559 AA; 62223 MW; FF485FDA78B40F20 CRC64;

Query Match 86.6%; Score 58; DB 2; Length 559;

Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
Db 201 PPAYPPPPVP 210

RESULT 9
Q8C220_MOUSE PRELIMINARY; ID 559 AA.
AC Q8C220;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Mus musculus B6-derived CD11 +ve dendritic cells cDNA, RIKEN full-length enriched library, clone:F730032N16 product:SH3-domain binding protein 2, full insert sequence.
DE protein 2, full insert sequence.
GN Name=Sh3bp2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; OC Muroidea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/350555500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I., Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S., Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J;
RA The FANTOM Consortium,
RT the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P., Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M., Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A., Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K., Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M., Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P., Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W., Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T., Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T., Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M., Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y., Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H., Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M., Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T., Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 PH domain.
DR EMBL; AK089455; BAC40891.1; -; mRNA.
DR Ensembl; ENSMUSG00000054520; Mus musculus.
DR MGI; MGI:1346349; Sh3bp2.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR001849; PH.
DR InterPro; IPR011993; PH_type.
DR InterPro; IPR000980; SH2.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00017; SH2; 1.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS50003; PH DOMAIN; 1.
DR PROSITE; PS50001; SH2; 1.
KW SH2 domain.
SQ SEQUENCE 559 AA; 62218 MW; 62248F9AF849E79 CRC64;

Query Match 86.6%; Score 58; DB 2; Length 559;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
Db 201 PPAYPPPPVP 210

RESULT 10
3BP2_HUMAN
ID 3BP2_HUMAN STANDARD; PRT; 561 AA.
AC P78314; O00500; O15373; P78315;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE SH3 domain-binding protein 2 (3BP-2).
GN Name=SH3BP2; Synonyms=3BP2; ORFNames=RES4-23;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM LONG).
RX MEDLINE=97446134; PubMed=9299232; DOI=10.1006/geno.1997.4849;
RA Bell S.M., Shaw M., Jou Y.-S., Myers R.M., Knowles M.A.;

RT "Identification and characterization of the human homologue of SH3BP2,
RT an SH3 binding domain protein within a common region of deletion at
RT 4p16.3 involved in bladder cancer."
RL Genomics 44:163-170(1997).
RN [2]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORMS LONG AND SHORT), AND TISSUE
RP SPECIFICITY.
RC TISSUE=Brain;
RX MEDLINE=98403881; PubMed=9734812;
RA Hadano S., Ishida Y., Ikeda J.-E.;
RT "The primary structure and genomic organization of five novel
RT transcripts located close to the Huntington's disease gene on human
RT chromosome 4p16.3";
RL DNA Res. 5:177-186(1998).
RN [3]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM LONG).
RC TISSUE=Tonsil;
RA Gokemeijer J., Deligiannidis K.E., Ligris K., Ernst T.J.;
RT "3BP2 binds to phosphatidylinositols; linking the hemopoietic tyrosine
RT kinase c-FES to the cytoplasmic membrane in a phosphorylation
RT dependent mechanism";
RL Blood 88:473A-473A(1996).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM LONG).
RC TISSUE=Cervix;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP SPLICE ISOFORM(S) THAT ARE POTENTIAL NMD TARGET(S).
RX PubMed=14759258; DOI=10.1186/gb-2004-5-2-r8;
RA Hillman R.T., Green R.E., Brenner S.E.;
RT "An unappreciated role for RNA surveillance.";
RL Genome Biol. 5:RESEARCH008.1-RESEARCH008.16(2004).
RN [6]
RP VARIANTS CRBM GLN-415; PRO-415; ARG-418; HIS-418; LEU-418; ARG-420 AND
RP GLU-420.
RX MEDLINE=21275962; PubMed=11381256; DOI=10.1038/88832;
RA Ueki Y., Tiziani V., Santanna C., Fukai N., Maulik C., Garfinkle J.,
RA Ninomiya C., doAmaral C., Peters H., Habal M., Rhee-Morris L.,
RA Doss J.B., Kreiborg S., Olsen B.R., Reichenberger E.;
RT "Mutations in the gene encoding c-Abl-binding protein SH3BP2 cause
RT cherubism.";
RL Nat. Genet. 28:125-126(2001).
RN [7]
RP VARIANT CRBM ARG-420.
RX PubMed=12900899; DOI=10.1002/ajmg.a.20226;
RA Lo B., Faiyaz-Ul-Haque M., Kennedy S., Aviv R., Tsui L.C., Teebi A.S.;
RT "Novel mutation in the gene encoding c-Abl-binding protein SH3BP2
RT causes cherubism";
RL Am. J. Med. Genet. 121:37-40(2003).
RN [8]
RP VARIANT CRBM ARG-418.
RX PubMed=14577811; DOI=10.1597/1545-1569(2003)040<0632:AMMITS>2.0.CO;2;
RA Imai Y., Kanno K., Moriya T., Kayano S., Seino H., Matsubara Y.,

RA Yamada A.;
RT "A missense mutation in the SH3BP2 gene on chromosome 4p16.3 found in
RT a case of nonfamilial cherubism.";
RL Cleft Palate Craniofac. J. 40:632-638(2003).
CC -|- FUNCTION: Binds differentially to the SH3 domains of certain
CC proteins of signal transduction pathways. Binds to
CC phosphatidylinositols; linking the hemopoietic tyrosine kinase fes
CC to the cytoplasmic membrane in a phosphorylation dependent
CC mechanism.
CC -|- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=P78314-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=P78314-2; Sequence=VSP_004085, VSP_004086;
CC Note=May be produced at very low levels due to a premature stop
CC codon in the mRNA, leading to nonsense-mediated mRNA decay;
CC -|- TISSUE SPECIFICITY: Expressed in a variety of tissues including
CC lung, liver, skeletal muscle, kidney and pancreas.
CC -|- DISEASE: Defects in SH3BP2 are the cause of cherubism (CRBM)
CC [MIM:118400]. CRBM is an autosomal dominant inherited syndrome
CC characterized by excessive bone degradation of the upper and lower
CC jaws, which often begins around three years of age. It is followed
CC by development of fibrous tissue masses, which causes a
CC characteristic facial swelling.
CC -|- SIMILARITY: Contains 1 PH domain.
CC -|- SIMILARITY: Contains 1 SH2 domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; U56386; AAB72034.1; -; mRNA.
CC EMBL; AB000462; BAA19119.1; -; mRNA.
CC EMBL; AB000463; BAA19120.1; -; mRNA.
CC EMBL; AF000936; AAB59973.1; -; mRNA.
CC EMBL; BC022996; AAH22996.1; -; mRNA.
CC HSSP; O08967; 1FW.
CC Ensembl; ENSG00000087266; Homo sapiens.
CC HGNC; HGNC:10825; SH3BP2.
CC H-InvDB; HIX0004037; -.
CC MIM; 602104; -.
CC MIM; 118400; -.
CC GO; GO:0005070; F:SH3/SH2 adaptor activity; TAS.
CC GO; GO:0007165; P:signal transduction; TAS.
CC InterPro; IPR001849; PH.
CC InterPro; IPR011993; PH_type.
CC InterPro; IPR000980; SH2.
CC Pfam; PF00169; PH; 1.
CC Pfam; PF00017; SH2; 1.
CC ProDom; PD000093; SH2; 1.
CC SMART; SM00233; PH; 1.
CC SMART; SM00252; SH2; 1.
CC PROSITE; PS50003; PH_DOMAIN; 1.
CC PROSITE; PS50001; SH2; 1.
KW Alternative splicing; Disease mutation; SH2 domain; SH3-binding.
FT DOMAIN 26 130
FT DOMAIN 457 555
FT MOTIF 201 210 SH3-binding.
FT COMPBIAS 205 212 Poly-Pro.
FT COMPBIAS 236 240 Poly-Pro.
FT VARSPPLIC 81 97 VMRAAETTSNNVFPFK -> QRPQPAQALSQTEAGP
FT (in isoform Short).
FT /FTId=VSP_004085.
FT Missing (in isoform Short).
FT /FTId=VSP_004086.
FT R -> P (in CRBM).
FT /FTId=VAR_013257.
FT R -> Q (in CRBM).
FT /FTId=VAR_013258.
FT P -> H (in CRBM).

```
FT          VARIANT      418      418      /FTid=VAR_013259.
FT          P -> L (in CRBM).
FT          /FTid=VAR_013260.
FT          VARIANT      418      418      P -> R (in CRBM).
FT          /FTid=VAR_013261.
FT          G -> E (in CRBM).
FT          /FTid=VAR_013262.
FT          VARIANT      420      420      G -> R (in CRBM).
FT          /FTid=VAR_013263.
FT          CONFLICT      27      27      V -> L (in Ref. 3).
FT          CONFLICT      224      224      H -> N (in Ref. 3).
FT          CONFLICT      249      249      L -> R (in Ref. 3).
FT          CONFLICT      251      251      A -> P (in Ref. 3).
SQ          SEQUENCE      561 AA; 62244 MW; 69E6846A4F6D8F15 CRC64;

Query Match      86.6%; Score 58; DB 1; Length 561;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY          1 PPPYPPPPPIP 10
Db          201 PPAYPPPPVP 210

RESULT 11
Q7PVW3 ANOGA
ID Q7PVW3_ANOGA PRELIMINARY; PRT; 1166 AA.
AC Q7PVW3;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000016732 (Fragment).
GN ORFNames=ENSANGG00000014243;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated (By
CC similarity).
CC -1- SIMILARITY: Contains 1 PH domain.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -1- SIMILARITY: Contains 1 PDZ (DHR) domain.
DR EMBL; AAB01008984; EAA15086.2; -; Genomic_DNA.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR001478; PDZ.
DR InterPro; IPR001849; PH.
DR InterPro; IPR001660; SAM.
DR Pfam; PF00595; PDZ; 1.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00536; SAM; 1; 1.
DR SMART; SM00228; PDZ; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS0106; PDZ; 1.
DR PROSITE; PS0003; PH DOMAIN; 1.
DR PROSITE; PS0105; SAM DOMAIN; 1.
FT NON TER      1166      1166
SQ          SEQUENCE      1166 AA; 128485 MW; AC823199F0AF1484 CRC64;
```

```
Query Match      86.6%; Score 58; DB 2; Length 1166;
Best Local Similarity 80.0%; Pred. No. 2.6e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY          1 PPPYPPPPPIP 10
Db          976 PPPLPPPPVP 985

RESULT 12
Q6Z1Z8 ORYSA
ID Q6Z1Z8_ORYSA PRELIMINARY; PRT; 1428 AA.
AC Q6Z1Z8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative PWWP domain protein.
GN Name=B1147B12.7-1;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Katayose Y.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005406; BAD03546.1; -; Genomic_DNA.
DR Gramene; Q6Z1Z8; -.
DR InterPro; IPR000313; PWWP.
DR InterPro; IPR006569; RPR.
DR Pfam; PF00855; PWWP; 1.
DR SMART; SM00293; PWWP; 1.
DR SMART; SM00582; RPR; 1.
DR PROSITE; PS0812; PWWP; 1.
KW Nucleotide-binding.
SQ          SEQUENCE      1428 AA; 155202 MW; 03FB9D3E6CA7FB65 CRC64;

Query Match      86.6%; Score 58; DB 2; Length 1428;
Best Local Similarity 80.0%; Pred. No. 3.2e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY          1 PPPYPPPPPIP 10
Db          1186 PPPIPPPPVP 1195

RESULT 13
Q00487_9CNID
ID Q00487_9CNID PRELIMINARY; PRT; 172 AA.
AC Q00487;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Mini-collagen (Fragment).
GN Name=N-COL 4;
OS Hydra sp.
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Anthomedusae;
OC Hydridae; Hydra.
OX NCBI_TaxID=6086;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SF1;
RX MEDLINE=92064646; PubMed=1955459; DOI=10.1083/jcb.115.4.1159;
RA Kurz E.M., Holstein T.W., Petri B.M., Engel J., David C.N.;
RT "Mini-collagens in hydra nematocytes.";
RL J. Cell Biol. 115:1159-1169 (1991).
DR EMBL; X61048; CAA43382.1; -; mRNA.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
```

DR InterPro; IPR008160; Collagen.
DR InterPro; IPR000976; Wilms tumour.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00049; WILMSTUMOUR.
KW Collagen.
FT NON TER 172 172
SQ SEQUENCE 172 AA; 16626 MW; F83AAD4369CFEB4F CRC64;

Query Match 85.1%; Score 57; DB 2; Length 172;
Best Local Similarity 90.0%; Pred. No. 46;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 57 PPPPPPPPPPIP 66

RESULT 14
Q6AVF5 ORYSA
ID Q6AVF5_ORYSA PRELIMINARY; PRT; 222 AA.
AC Q6AVF5;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Expressed protein.
GN Name=OSJNBb0058G04.9;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Taitrin T., Kim M.M., Bera J.J., Jin S.S.,
RA Padrosh D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblyum T.V.,
RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RT "Oryza sativa chromosome 3 BAC OSJNBb0058G04 genomic sequence."
RL Submitted (NOV-2001) to the EMBL/GenBank/DBSJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Buell R.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AC103551; AAT78807.1; -; Genomic_DNA.
DR Gramene; Q6AVF5; -.
DR InterPro; IPR008889; VQ.
DR Pfam; PF05678; VQ; 1.
SQ SEQUENCE 222 AA; 22873 MW; 08EFB6375A5F9E19 CRC64;

Query Match 85.1%; Score 57; DB 2; Length 222;
Best Local Similarity 80.0%; Pred. No. 60;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 7 PPHPPPPPPAP 16

RESULT 15
Q5AX03 EMENI
ID Q5AX03_EMENI PRELIMINARY; PRT; 231 AA.
AC Q5AX03;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein.
GN ORFNames=AN7177.2;
OS Aspergillus nidulans FGSC A4.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=227321;

RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FGSC A4;
RA Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavkiy L.,
RA Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArellano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA Roman J., Schauer S., Schupback R., Seaman S., Severy P., Smirnov S.,
RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA Lander E.;
RT "Genome Sequence of Aspergillus nidulans."
RL Submitted (JAN-2004) to the EMBL/GenBank/DBSJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBSJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACD01000122; EAA61429.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 231 AA; 25893 MW; CBF091B4D16D6941 CRC64;

Query Match 85.1%; Score 57; DB 2; Length 231;
Best Local Similarity 80.0%; Pred. No. 63;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYPPPPPIP 10
Db 88 PPSYPPPPPLP 97

RESULT 16
ATF5 HUMAN
ID ATF5_HUMAN STANDARD; PRT; 282 AA.
AC Q9Y2D1; Q9BSA1; Q9UNQ3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Cyclic-AMP-dependent transcription factor ATF-5 (Activating
DE transcription factor 5) (Transcription factor ATFx).
GN Name=ATF5; Synonyms=ATFX;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA], AND VARIANT PRO-121.
RC TISSUE=Brain;
RX MEDLINE=20558615; PubMed=11087824; DOI=10.1073/pnas.240452197;
RA White J.H., McIlhinney R.A.J., Wise A., Ciruela F., Chan W.-Y.,
RA Emson P.C., Billinton A., Marshall F.H.;
RT "The GABAB receptor interacts directly with the related transcription
RT factors CREB2 and ATFx."
RL Proc. Natl. Acad. Sci. U.S.A. 97:13967-13972(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC TISSUE=Hepatoblastoma;
RX PubMed=15221005; DOI=10.1038/sj.onc.1207782;
RA Yamada S., Ohira M., Horie H., Ando K., Takayasu H., Suzuki Y.,
RA Sugano S., Hirata T., Goto T., Matsunaga T., Hiyama E., Hayashi Y.,

RA Ando H., Suita S., Kaneko M., Sasaki F., Hashizume K., Ohnuma N.,
RA Nakagawara A.;
RT "Expression profiling and differential screening between
RT hepatoblastomas and the corresponding normal livers: identification of
RT high expression of the PLK1 oncogene as a poor-prognostic indicator of
RT hepatoblastomas.";
RL Oncogene 23:5901-5911(2004).
RN [3]
RP NUCLEOTIDE SEQUENCE [MRNA], AND VARIANT PRO-121.
RA Kohroki J., Tanaka K.;
RT "cDNA clone encoding leucine-zipper protein.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Touchman J.W., Green E.D., Dickson M.C.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE [MRNA] OF 161-282.
RX MEDLINE=99303793; PubMed=10373550;
RA Pati D., Meistrich M.L., Plon S.E.;
RT "Human Cdc34 and Rad6B ubiquitin-conjugating enzymes target repressors
RT of cyclic AMP-induced transcription for proteolysis.";
RL Mol. Cell. Biol. 19:5001-5013(1999).
RN [6]
RP INTERACTION WITH CCND3, FUNCTION, AND SUBCELLULAR LOCATION.
RX PubMed=15358120; DOI=10.1016/j.bbrc.2004.07.053;
RA Liu W., Sun M., Jiang J., Shen X., Sun Q., Liu W., Shen H., Gu J.;
RT "Cyclin D3 interacts with human activating transcription factor 5 and
RT potentiates its transcription activity.";
RL Biochem. Biophys. Res. Commun. 321:954-960(2004).
CC -!- FUNCTION: Transcriptional activator which binds the cAMP response
CC element (CRE) (consensus: 5'-GTGACGT[AC][AG]-3'), a sequence
CC present in many viral and cellular promoters and blocks the
CC differentiation of neuroprogenitor cells into neurons. Its
CC transcriptional activity is enhanced by CCND3 and slightly
CC inhibited by CDK4.
CC -!- SUBUNIT: Binds DNA as a dimer. Interacts with PTP4A1/PRL-1 (By
CC similarity). Interacts with CCND3, but not with CCND1 or CCND2.
CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.
CC -!- SIMILARITY: Belongs to the bZIP family.
CC -!- SIMILARITY: Contains 1 bZIP domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AF305687; AAG22558.1; -; mRNA.
CC EMBL; AB073613; BAD38650.1; -; mRNA.
CC EMBL; AB021663; BAA78477.2; -; mRNA.
CC EMBL; BC005174; AAH05174.1; -; mRNA.
CC EMBL; AF101388; AAD28370.1; -; mRNA.

DR HSSP; P18848; 1Ci6.
DR TRANSPAC; T04877; -.
DR Ensembl; ENSG00000169136; Homo sapiens.
DR HGNC; HGNC:790; ATF5.
DR MIM; 606398; -.
DR GO; GO:0003702; F:RNA polymerase II transcription factor acti. . .; TAS.
DR GO; GO:0003714; F:transcription corepressor activity; TAS.
DR GO; GO:0000074; P:regulation of cell cycle; TAS.
DR GO; GO:0006357; P:regulation of transcription from RNA polyme. . .; TAS.
DR InterPro; IPR011616; bZIP 1.
DR InterPro; IPR004827; TF_bZIP.
DR Pfam; PF00170; bZIP 1; 1.
DR SMART; SM00338; BRLZ; 1.
DR PROSITE; PS02017; bZIP; 1.
DR PROSITE; PS00036; bZIP_BASIC; FALSE_NEG.
KW Activator; DNA-binding; Multigene family; Nuclear protein;
KW Polymorphism; Transcription; Transcription regulation.
FT DOMAIN 236 250 Leucine-zipper (Probable).
FT DNA_BIND 210 230 Basic motif.
FT REGION 119 217 Interaction with PTP4A1 (By similarity).
FT COMPIAS 123 139 Poly-Pro.
FT COMPIAS 186 194 Poly-Pro.
FT VARIANT 121 121 L -> P (in dbSNP:283526).
FT /FTID=VAR 022786.
FT LLA -> RHE (in Ref. 5).
SQ SEQUENCE 282 AA; 30690 MW; B6D4CB6F631655B1 CRC64;

Query Match 85.1%; Score 57; DB 1; Length 282;
Best Local Similarity 80.0%; Pred. No. 77;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPIP 10
DB 127 PPPLPPPPPLP 136

RESULT 17
Q8IW12 HUMAN
ID Q8IW12_HUMAN PRELIMINARY; PRT; 434 AA.
AC Q8IW12;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Transcription factor AP-2 epsilon (Activating enhancer binding protein
DE 2 epsilon).
GN Name=TFAP2E;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Uterus;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Uterus;
RA Director MGC Project;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC041175; AAH41175.1; -; mRNA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR004979; TF_AP2.
DR Pfam; PF03299; TF_AP-2; 1.
DR PRINTS; PR01748; AP2TNSCPFCT.
SQ SEQUENCE 434 AA; 45308 MW; 6209DD5D07B8872C CRC64;

Query Match 85.1%; Score 57; DB 2; Length 434;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
||||| |||:
Db 48 PPPYPQPLP 57

RESULT 18
Q6VUC0 HUMAN
ID Q6VUC0_HUMAN PRELIMINARY; PRT; 442 AA.
AC Q6VUC0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Transcription factor AP-2 epsilon (Activating enhancer binding protein 2 epsilon).
DE 2 epsilon).
GN Name=TFAP2E; ORFNames=RP5-983H21.2-001;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22999112; PubMed=14636996; DOI=10.1016/S0378-1119(03)00840-0;
RA Tummala R., Romano R.A., Fuchs E., Sinha S.;
RT "Molecular cloning and characterization of AP-2 epsilon, a fifth member of the AP-2 family.";
RL Gene 321:93-102(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Heath P.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY326454; AAQ91614.1; -; mRNA.
DR EMBL; AL157951; CAI23520.1; -; Genomic DNA.
DR EMBL; AC004865; CAI23520.1; JOINED; Genomic_DNA.
DR Ensembl; ENSG00000116819; Homo sapiens.
DR HGNC; HGNC:30774; TFAP2E.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR004979; TF_AP2.
DR Pfam; PF03299; TF_AP-2; 1.
DR PRINTS; PR01748; AP2TNSCPFCT.
SQ SEQUENCE 442 AA; 46212 MW; 823209DA7C0EC527 CRC64;

Query Match 85.1%; Score 57; DB 2; Length 442;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
||||| |||:
Db 56 PPPYPQPLP 65

RESULT 19
Q61WU8 CAEBR
ID Q61WU8_CAEBR PRELIMINARY; PRT; 521 AA.
AC Q61WU8;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG04249.
GN Name=CBG04249;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG The C.briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.
CC EMBL; CAAC0100017; CAE60609.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 521 AA; 59983 MW; 40D732BAD6BAAC80 CRC64;

Query Match 85.1%; Score 57; DB 2; Length 521;
Best Local Similarity 90.0%; Pred. No. 1.5e+02;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
||||| |||:
Db 188 PPPPPPPPIP 197

RESULT 20
Q5SIV4 THET8
ID Q5SIV4_THET8 PRELIMINARY; PRT; 590 AA.
AC Q5SIV4;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE DNA primase.
GN OrderedLocusNames=TTHA1260;
OS Thermus thermophilus (strain HB8 / ATCC 27634 / DSM 579).
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermaceae;
OC Thermus.
OX NCBI_TaxID=300852;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HB8;
RA Masui R., Kurokawa K., Nakagawa N., Tokunaga F., Koyama Y., Shibata T., Oshima T., Yokoyama S., Yasunaga T., Kuramitsu S.;
RT "Complete genome sequence of Thermus thermophilus HB8.";
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP008226; BAD71083.1; -; Genomic_DNA.
DR GO; GO:0005658; C:alpha DNA polymerase:primase complex; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003896; F:DNA primase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006304; P:DNA modification; IEA.
DR GO; GO:0006260; P:DNA replication; IEA.
DR GO; GO:0006269; P:DNA replication, synthesis of RNA primer; IEA.
DR InterPro; IPR006295; DNA_primase.
DR InterPro; IPR006171; Toprim_dom.
DR InterPro; IPR006647; Toprim_primease.
DR InterPro; IPR006154; Toprim_sub.
DR InterPro; IPR002694; Znf_CHC2.
DR Pfam; PF01751; Toprim; 1.
DR Pfam; PF01807; zf-CHC2; 1.
DR ProDom; PD002276; Toprim_primease; 1.
DR ProDom; PD002988; Znf_CHC2; 1.
DR SMART; SM00493; TOPRIM; 1.
DR SMART; SM00400; Znf_CHCC; 1.
DR TIGRFAMs; TIGR01391; dnaG; 1.


```

SQ SEQUENCE 802 AA; 85844 MW; 592BB975EE20F77F CRC64;

Query Match 85.1%; Score 57; DB 1; Length
Best Local Similarity 80.0%; Pred. No. 2.3e+02;
Matches 8; Conservative 1; Mismatches 1; Index

```

```

RESULT 22
O48682_ARATH PRELIMINARY;      PRT;      820 AA.
AC O48682;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE F3I6.8 protein.
GN Name=F3I6.8;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Federspiel N.A., Palm C.J., Conway A.B., Kurtz D.B., Conway A.R.,
RA Au M., Araujo R., Buehler E., Dewar K., Feng J., Kim C., Li Y.,
RA O'J O., Osborne B.I., Shinn P., Sun H., Toriumi M., Vysotskaia V.S.,
RA Yu G., Ecker J., Theologis A., Davis R.W.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC002396; AAC00575.1; -; Genomic_DNA.
DR PIR; T00645; T00645.

```

RESULT 23	FRU_DROME	FRU_DROME	STANDARD;	PRT;	955 AA.
ID	AC	Q8IN81;	O44708;	P91619;	P91619; Q24004; Q8IN80; Q8IN82; Q8IN83;
AC	AC	Q9GU18;	Q9GU19;	Q9GU20;	Q9GU21; Q9GU22; Q9VE64; Q9VE65; Q9VE66;
AC	AC	Q9VE67;			
DT	DT	10-OCT-2003	(Rel. 42,	Created)	
DT	DT	10-OCT-2003	(Rel. 42,	Last sequence update)	
DT	DT	13-SEP-2005	(Rel. 48,	Last annotation update)	
DE	DE	Sex determination	protein	fruitless.	
GN	GN	Name=fru;	Synonyms=BtB-VI;	ORFNames=CG14307;	
OS	OS	Drosophila melanogaster	(Fruit fly).		
OC	OC	Eukaryota;	Metazoa;	Arthropoda;	Insecta; Pterygota;
OC	OC	Neoptera;	Endopterygota;	Diptera;	Brachycera; Muscomorpha;
OC	OC	Ephydroidea;	Drosophilidae;	Drosophila.	
OX	OX	NCBI_TaxID=7227;			
RN	RN	{1}			
RP	RP	NUCLEOTIDE SEQUENCE (ISOFORMS	MALE-I AND FEMALE-I),	FUNCTION,	TISSUE
RP	RP	SPECIFICITY, AND INTERACTION	WITH TRA.		
RC	RC	STRAIN=Oregon-R;			
RX	RX	MEDLINE=97133215;	PubMed=8978612;	DOI=10.1016/S0092-8674(00)81802-4;	
RA	RA	Ryner L.C., Goodwin S.F.,	Castrillon D.H., Anand A.,	Villella A.,	
RA	RA	Baker B.S., Hall J.C.,	Taylor B.J., Wasserman S.A.;		

RT "Control of male sexual behavior and sexual orientation in *Drosophila*
RT by the fruitless gene.";
RL Cell 87:1079-1089(1996).
RN [2]

RP NUCLEOTIDE SEQUENCE (ISOFORM FEMALE-A), FUNCTION, AND TISSUE
RP SPECIFICITY.
RC STRAIN=Canton-S; TISSUE=Head;
RX MEDLINE=96382528; PubMed=8790392; DOI=10.1073/pnas.93.18.9687;
RA Ito H., Fujitani K., Usui K., Shimizu-Nishikawa K., Tanaka S.,
RA Yamamoto D.;
RT "Sexual orientation in *Drosophila* is altered by the satori mutation in
RT the sex-determination gene fruitless that encodes a zinc finger
RT protein with a BTB domain.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:9687-9692(1996).
RN [3]

RP NUCLEOTIDE SEQUENCE (ISOFORMS MALE-A; FEMALE-A; MALE-B; FEMALE-B;
RP TYPE-C; TYPE-D; MALE-E AND FEMALE-E), FUNCTION, TISSUE SPECIFICITY,
RP AND INTERACTION WITH TRA.
RC STRAIN=Canton-S; TISSUE=Head;
RX MEDLINE=20394292; PubMed=10934470; DOI=10.1038/35019537;
RA Usui-Aoki K., Ito H., Ui-Tei K., Takahashi K., Lukacsovich T.,
RA Awano W., Nakata H., Piao Z.F., Nilsson E.E., Tomida J.-Y.,
RA Yamamoto D.;
RT "Formation of the male-specific muscle in female *Drosophila* by ectopic
RT fruitless expression.";
RL Nat. Cell Biol. 2:500-506(2000).
RN [4]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner K., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
RN [5]

RP GENOME REANNOTATION, AND ALTERNATIVE SPLICING.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradscky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,

Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
Lewis S.E.;
"Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review";
Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
[6]
NUCLEOTIDE SEQUENCE OF 104-218.
MEDLINE=95024186; PubMed=7938017;
Zollman S., Godt D., Prive G.G., Couderc J.-L., Laski F.A.;
"The BTB domain, found primarily in zinc finger proteins, defines an
evolutionarily conserved family that includes several developmentally
regulated genes in Drosophila.";
Proc. Natl. Acad. Sci. U.S.A. 91:10717-10721(1994).
[7]
FUNCTION.
MEDLINE=98078701; PubMed=9418892;
Heinrichs V., Ryner L.C., Baker B.S.;
"Regulation of sex-specific selection of fruitless 5' splice sites by
transformer and transformer-2.";
Mol. Cell. Biol. 18:450-458(1998).
[8]
FUNCTION.
STRAIN=Canton-S;
MEDLINE=22323098; PubMed=12435630; DOI=10.1101/gad.1010302;
Dauwalder B., Tsujimoto S., Moss J., Mattox W.;
"The Drosophila takeout gene is regulated by the somatic sex-
determination pathway and affects male courtship behavior.";
Genes Dev. 16:2879-2892(2002).
-!- FUNCTION: Probably acts as a transcriptional regulator. Part of
the somatic sex determination hierarchy; sex determination genes
transformer (tra) and transformer-2 (tra-2) switch fru splicing
from the male-specific pattern to the female-specific pattern
through activation of the female-specific fru 5' splice site.
Vital for the development of males and females. Controls the
development of the male specific abdominal muscle of Lawrence.
Plays a role in male courtship behavior and sexual orientation.
Enhances male-specific expression of takeout in brain-associated
fat body.
-!- INTERACTION:
Q9V4B8:CG10323; NbExp=1; IntAct=EBI-196831, EBI-145217;
-!- SUBCELLULAR LOCATION: Nuclear (Potential).
-!- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=10;
Comment=Sex specific splicing is controlled by tra and tra-2.
Sex-specific repression of the translation of fru mRNA might be
mediated by the binding of Tra to the fru mRNA in females;
Name=Male-A; Synonyms=E;
IsoId=Q8IN81-1; Sequence=Displayed;
Name=Female-A; Synonyms=C;
IsoId=Q8IN81-4; Sequence=VSP_050497;
Name=Female-B; Synonyms=F;
IsoId=Q8IN81-7; Sequence=VSP_050497, VSP_050501, VSP_050505;
Note=Ref.3 (AAG28589) sequence is in conflict in position:
650:T->A;
Name=Female-E; Synonyms=H;
IsoId=Q8IN81-10; Sequence=VSP_050497, VSP_050500, VSP_050506;
Name=Female-I;
IsoId=Q8IN81-3; Sequence=VSP_050497, VSP_050498, VSP_050501,
VSP_050505;
Name=Male-B; Synonyms=B;
IsoId=Q8IN81-6; Sequence=VSP_050501, VSP_050505;
Name=Male-E; Synonyms=G;
IsoId=Q8IN81-12; Sequence=VSP_050500, VSP_050506;
Name=Male-I;
IsoId=Q8IN81-2; Sequence=VSP_050498, VSP_050501, VSP_050505;
Name=Type-C; Synonyms=D;
IsoId=Q8IN81-8; Sequence=VSP_050497, VSP_050499, VSP_050504;
Note=Ref.3 (AAG28590) sequence is in conflict in positions:
539:S->P, 617:H->R;
Name=Type-D; Synonyms=A;

ISOId=Q8IN81-9; Sequence=VSP_050497, VSP_050502, VSP_050503;
-!- TISSUE SPECIFICITY: Expressed in parts of the adult male brain
associated with the courtship song and steps of the male
courtship. Also expressed in the larval and pupal male mushroom
body and optic lobe. Expressed in pupal female optic lobe.
-!- MISCELLANEOUS: Mutant males exhibit bisexual behavior; they court
females but are behaviorally sterile so fail to mate and they
exhibit vigorous courtship with other fru mutant males.
-!- SIMILARITY: Contains 1 BTB (POZ) domain.
-!- SIMILARITY: Contains 1 C2H2-type zinc finger.

This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.

EMBL; AF039231; AAB96677.1; -; mRNA.
EMBL; U72492; AAB92662.1; -; mRNA.
EMBL; D84437; BAA12663.1; -; mRNA.
EMBL; D84438; BAA12664.1; -; mRNA.
EMBL; AF220176; AAG28587.1; -; mRNA.
EMBL; AF220177; AAG28588.1; -; mRNA.
EMBL; AF220178; AAG28589.1; -; mRNA.
EMBL; AF220179; AAG28590.1; -; mRNA.
EMBL; AF220180; AAG28591.1; -; mRNA.
EMBL; AF220181; AAG28592.1; -; mRNA.
EMBL; AE003722; AAF55562.2; -; Genomic DNA.
EMBL; AE003722; AAF55563.2; -; Genomic DNA.
EMBL; AE003722; AAF55564.2; -; Genomic DNA.
EMBL; AE003722; AAF55565.2; -; Genomic DNA.
EMBL; AE003722; AAN13774.1; -; Genomic DNA.
EMBL; AE003722; AAN13775.1; -; Genomic DNA.

Query Match 85.1%; Score 57; DB 1; Length 955;
Best Local Similarity 80.0%; Pred. No. 2.7e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 PPYP PPPPIP 10
||| |||||
Db 78 PPPL PPPPLP 87

RESULT 24
Q96JH1_HUMAN
ID Q96JH1_HUMAN PRELIMINARY; PRT; 1134 AA.
AC Q96JH1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE KIAA1856 protein (Fragment).
GN Name=KIAA1856;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=21245130; PubMed=11347906;
RA Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XX.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 8:85-95(2001).
DR EMBL; AB058759; BAB47485.1; -; mRNA.
DR Ensembl; ENSG00000182095; Homo sapiens.
DR InterPro; IPR000637; A+T_hook.
DR PRINTS; PR00929; ATHOOK.
FT NON_TER 1
SQ SEQUENCE 1134 AA; 122299 MW; 1F1AA902C19F2EC2 CRC64;

Query Match 85.1%; Score 57; DB 2; Length 1134;
Best Local Similarity 80.0%; Pred. No. 3.2e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
Db 969 PPPLPPPPPLP 978
||| |||||:|

RESULT 25
Q6PLP6_CHLRE PRELIMINARY; PRT; 1146 AA.
AC Q6PLP6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cell wall protein GP2 (Fragment).
GN Name=GP2;
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae;
OC Chlamydomonadales; Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C9;
RA Ferris P.J., Small L., Goodenough U.W.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AY596305; AAT02521.1; -; mRNA.
DR InterPro; IPR002951; Atrophin.
DR PRINTS; PR01222; ATROPHIN.
FT NON TER 1
SQ SEQUENCE 1146 AA; 120311 MW; 749663DA78E316E3 CRC64;

Query Match 85.1%; Score 57; DB 2; Length 1146;
Best Local Similarity 80.0%; Pred. No. 3.3e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
Db 909 PPPSPPPPPVP 918
||| |||||:

RESULT 26
Q4WGB0_ASPFU PRELIMINARY; PRT; 1273 AA.
ID Q4WGB0_ASPFU
AC Q4WGB0;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=Afu7G04400;
OS Aspergillus fumigatus Af293.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=330879;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Af293;
RA Nierman W., Pain A., Anderson M.J., Wortman J., Kim H.Stanley.,
RA Arroya J., Berriman M., Abe K., Archer D.B., Bermejo C., Bennett J.,
RA Bowyer P., Chen D., Collins M., Coulson R., Davies R., Dyer P.S.,
RA Farman M., Fedorova N., Fedorova N., Feldblyum T.V., Fischer R.,
RA Foerster N., Fraser A., Garcia J.L., Garcia M.J., Goble A.,
RA Goldman G.H., Gomi K., Griffith-Jones S., Gwilliam R., Haas B.,
RA Haas H., Harris D., Horiuchi H., Huang J., Humphrey S., Jimenez J.,
RA Keller N., Khouri H., Kitamoto K., Kobayashi T., Kulkarni R.,
RA Kumagai T., Lafton A., Latge J.-P., Li W., Lord A., Lu C.,
RA Majoros W.H., May G.S., Miller B.L., Mohamoud Y., Molina M., Monod M.,
RA Mouyna I., Mulligan S., Murphy L., O'Neil S., Paulsen I.,
RA Penalba M.A., Pertea M., Price C., Pritchard B.L., Quail M.A.,
RA Rabinowitsch E., Rawlins N., Rajandream M.-A., Reichard U.,
RA Renauld H., Robson G.D., Rodriguez de Cordoba S., Rodriguez-Pena J.M.,
RA Ronning C.M., Rutter S., Salzberg S.L., Sanchez M.,

RA Sanchez-Ferrero J.C., Saunders D., Seeger K., Squares R., Squares S.,
RA Takeuchi M., Tekala F., Turner G., Vazquez de Aldana C.R., Weidman J.,
RA White O., Woodward J., Yu J.-H., Fraser G., Galagan J.E., Asai K.,
RA Machida M., Hall N., Barrell B., Denning D.W.;
RT "Genomic sequence of the pathogenic and allergenic filamentous fungus
RT Aspergillus fumigatus.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBSJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBSJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHF0100009; EAL87031.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 1273 AA; 140074 MW; 9C80E6B6BDCF8B14 CRC64;

Query Match 85.1%; Score 57; DB 2; Length 1273;
Best Local Similarity 90.0%; Pred. No. 3.6e+02;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
Db 225 PPPPPPPPPPIP 234
||| |||||

RESULT 27
Q8MMQ2_DICDI PRELIMINARY; PRT; 1324 AA.
ID Q8MMQ2_DICDI
AC Q8MMQ2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein.
GN ORFNames=DDB0167272;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RX MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J.A.,
RA Bankier A.T., Dear P.H., Lehmann R., Baumgart C., Paira G.,
RA Abril J.F., Guigo R., Kumpf K., Tunggal B., Cox E., Quail M.A.,
RA Platzter M., Rosenthal A., Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
RL Nature 418:79-85(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RA Baumgart C.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBSJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RA Sugang R., Berriman M., Song J., Olsen R., Szafranski K., Xu Q.,
RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,
RA Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA Muzny D., Mourier T., Pain A., Lu M., Harper D., Lindsay R.,
RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA Wardroper A., Felder M., Thangavelu M., Johnson D., Knights A.,
RA Loulseged H., Mungall K., Oliver K., Price C., Quail M.A.,
RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,
RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
RA Shaalsky G., Schleicher M., Weinstock G., Rosenthal A., Cox E.C.,
RA Chisholm R.L., Gibbs R., Loomis W.F., Platzter M., Kay R.R.,
RA Williams J., Dear P.H., Noegel A.A., Barrell B., Kuepa A.;
RT "The genome of the social amoeba Dictyostelium discoideum.";
RL Nature 0:0-0(2005).

DR EMBL; AC117080; AAM45316.2; -; Genomic_DNA.
DR EMBL; AAFI01000030; EAL69520.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 1324 AA; 148579 MW; F5C922EB79DDAB40 CRC64;

Query Match 85.1%; Score 57; DB 2; Length 1324;
Best Local Similarity 90.0%; Pred. No. 3.8e+02;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPPPPPPIP 10
||| |||||
Db 688 PPPPPPPPIP 697

RESULT 28
Q7SC01 NEUCR PRELIMINARY; PRT; 1395 AA.
AC Q7SC01;
DT 01-MAR-2004 (TReMBLrel. 26, Created)
DT 01-MAR-2004 (TReMBLrel. 26, Last sequence update)
DT 01-FEB-2005 (TReMBLrel. 29, Last annotation update)
DE Predicted protein (Hypothetical protein B22K18.130).
GN Name=NCU09482.1; Synonyms=B22K18.130;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=OR74A;

RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA Elkins T., Engels P., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA Kothe G.O., Jedd G., Meves W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA Kamal M., Kamvysselis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
RA Krystofova S., Rasmussen C., Metzenberg R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmani S.A.,
RA DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.;
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL Nature 0:0-0(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA German Neurospora genome project;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.

DR EMBL; AABX01000145; EAA33970.1; -; Genomic_DNA.
DR EMBL; BX842597; CAE75735.1; -; Genomic_DNA.
DR HSSP; Q92831; 1JM4.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0007049; P:cell cycle; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001487; Bromodomain.
DR PROSITE; PS50014; BROMODOMAIN_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 1395 AA; 149956 MW; CB0D84322CDA8835 CRC64;

Query Match 85.1%; Score 57; DB 2; Length 1395;
Best Local Similarity 90.0%; Pred. No. 4e+02;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPPPPPPIP 10
||| |||||
Db 341 PPPPPPPPIP 350

RESULT 29
Q9VDD2 DROME PRELIMINARY; PRT; 1400 AA.
ID Q9VDD2_DROME PRELIMINARY;
AC Q9VDD2; Q8SZS7;

DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 10-MAY-2005 (TReMBLrel. 30, Last annotation update)
DE CG17299-PF, isoform F (LD222662p) (Loechnig isoform I).
GN Name=SNF4A-gamma; Synonyms=SNF4Agamma, loe; ORFNames=CG17299;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]

RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20196006; PubMed=107311132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426065; PubMed=12537568;

RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskaas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
RT melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).

RN [3] NUCLEOTIDE SEQUENCE.
RP MEDLINE=22426070; PubMed=12537573;
RX Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
a genomics perspective.";
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426069; PubMed=12537572;
RA Milara S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RG Berkeley Drosophila Genome Project;
RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
RA Yu C., Rubin G.;
RT "Drosophila melanogaster release 4 sequence.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE.
RG FlyBase;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
RN [7]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RN [8]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22344351; PubMed=12456644; DOI=10.1093/emboj/cdf636;
RA Tschaep J.A., Hammerschmid C., Muhlig-Versen M., Athenstaedt K.,
RA Daum G., Kretschmar D.;
RT "The neurodegeneration mutant lochrig interferes with cholesterol
homeostasis and Appl processing.";
RL EMBO J. 21:6367-6376(2002).
RN [9]
RP NUCLEOTIDE SEQUENCE.
RA Tschaep J.A.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003733; AAF55864.2; -; Genomic_DNA.
DR EMBL; AY070541; AAL48012.1; -; mRNA.
DR EMBL; AY166752; AAN85714.1; -; mRNA.
DR FlyBase; FBgn0025803; SNF4A-gamma.
DR InterPro; IPR000644; CBS.
DR Pfam; PF00571; CBS; 2.
DR SMART; SM00116; CBS; 4.
SQ SEQUENCE 1400 AA; 152381 MW; 411B93CC6B9EC7AF CRC64;

Query Match 85.1%; Score 57; DB 2; Length 1400;
Best Local Similarity 90.0%; Pred. No. 4e+02;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPIP 10
Db 280 PPPPPPPPIP 289

RESULT 30
Q752A6 ASHGO
ID Q752A6 ASHGO PRELIMINARY; PRT; 1918 AA.
AC Q752A6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE AFR669WP.
GN Name=AFR669W;
OS Ashbya gossypii (Yeast) (Eremothecium gossypii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Eremothecium.
OX NCBI_TaxID=33169;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 10895;
RX PubMed=15001715; DOI=10.1126/science.1095781;
RA Dietrich F.S., Voegeli S., Brachat S., Lerch A., Gates K., Steiner S.,
RA Mohr C., Poehlmann R., Luedi P., Choi S., Wing R.A., Flavien A.,
RA Gaffney T.D., Philippsen P.;
RT "The Ashbya gossypii genome as a tool for mapping the ancient
Saccharomyces cerevisiae genome.";
RL Science 304:304-307(2004).
DR EMBL; AE016819; AAS54041.1; -; Genomic_DNA.
DR SMR; Q752A6; 1369-1776.
DR AGD; AFR669W; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003779; F:actin binding; IEA.
DR GO; GO:0017048; F:Rho GTPase binding; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0030036; P:actin cytoskeleton organization and biogenesis; IEA.
DR GO; GO:0016043; P:cell organization and biogenesis; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR010472; Drf_FH3.
DR InterPro; IPR010473; Drf_GBD.
DR InterPro; IPR003104; FH2_actin_bd.
DR InterPro; IPR000976; Wilms_tumour.
DR Pfam; PF06367; Drf_FH3; 1.
DR Pfam; PF06371; Drf_GBD; 1.
DR Pfam; PF02181; FH2; 1.
DR PRINTS; PR00049; WILMSTUMOUR.
DR SMART; SM00498; FH2; 1.
KW Complete proteome.
SQ SEQUENCE 1918 AA; 213018 MW; 2E07CC6AE7B9B0A6 CRC64;

Query Match 85.1%; Score 57; DB 2; Length 1918;
Best Local Similarity 90.0%; Pred. No. 5.6e+02;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPIP 10
Db 1256 PPPPPPPPIP 1265

RESULT 31
Q5JM22 ORYSA
ID Q5JM22 ORYSA PRELIMINARY; PRT; 135 AA.
AC Q5JM22;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein P0702H08.26.
GN Name=P0702H08.26;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Katayose Y.,
RA Wu J., Niimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,

RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
RA Ikeno M., Itoh S., Itoh T., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,
RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagasaki H., Nakashima M., Nakama Y., Nakanichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tauji K.,
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA Yano M., Jiang J., Gojobori T.,
RT "The genome sequence and structure of rice chromosome 1.";
RL Nature 420:312-316(2002).
DR EMBL; AP003350; BAD87477.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 135 AA; 13727 MW; 72D1DF60F44E4E71 CRC64;

Query Match 83.6%; Score 56; DB 2; Length 135;
Best Local Similarity 80.0%; Pred. No. 46;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPIP 10
Db 35 PPPPPPPPPVP 44

RESULT 32
Q7YTN7 CAEEL
ID Q7YTN7 CAEEL PRELIMINARY; PRT; 170 AA.
AC Q7YTN7;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein F20G2.6a.
GN ORFNames=F20G2.6, F20G2.6a;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z79753; CAE17798.1; -; Genomic_DNA.
DR Ensembl; F20G2.6; Caenorhabditis elegans.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 170 AA; 18917 MW; 831A696032A97804 CRC64;

Query Match 83.6%; Score 56; DB 2; Length 170;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPYPPPP 8
Db 125 PPPYPPPP 132

RESULT 33
Q61FU6 CAEHR
ID Q61FU6 CAEHR PRELIMINARY; PRT; 173 AA.
AC Q61FU6;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG11492.
GN Name=CBG11492;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG The C.briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAC01000054; CAE66248.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 173 AA; 19188 MW; B53564996FDBF2F4 CRC64;

Query Match 83.6%; Score 56; DB 2; Length 173;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPYPPPP 8
Db 128 PPPYPPPP 135

RESULT 34
Q52GY4 CAEEL
ID Q52GY4 CAEEL PRELIMINARY; PRT; 218 AA.
AC Q52GY4;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein F20G2.6b.
GN ORFNames=F20G2.6B;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z79753; CAI91165.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 218 AA; 24544 MW; 9A9D014CAFC005DF CRC64;

Query Match 83.6%; Score 56; DB 2; Length 218;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPYPPPP 8
Db 173 PPPYPPPP 180

RESULT 35
Q64M77 ORYSA
ID Q64M77 ORYSA PRELIMINARY; PRT; 233 AA.
AC Q64M77;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein OSJNBa0004005.28.
GN Name=OSJNBa0004005.28;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Fujisawa M.;

RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 2, BAC
RT clone:OSJNBa004O05."
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP007253; BAD46728.1; -; Genomic_DNA.

DR Gramene; Q64M77; -
DR InterPro; IPR001012; UBX.
DR Pfam; PF00789; UBX; 1.
DR PROSITE; PS50033; UBX; 1.
KW Hypothetical protein.
SQ SEQUENCE 233 AA; 25299 MW; 3E0F3119708A4D7B CRC64;

Query Match 83.6%; Score 56; DB 2; Length 233;
Best Local Similarity 80.0%; Pred. No. 81;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYP PPPPIP 10
||| |||||:
Db 82 PPPPP PPPPVP 91

RESULT 36

Q4P543 USTMA PRELIMINARY; PRT; 247 AA.

AC Q4P543;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=UM04770.1;
OS Ustilago maydis 521.
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OX NCBI_TaxID=237631;
RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=521;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
RA Bayul T., Blitshsteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA Gnirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamvysseilis M., Karlsson E.,
RA Kells C., Kieu A., Kianer P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., Mcdonough S., Mcghee T., Meldrim J., Meneus L.,
RA Mesirov J., Mihalev A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., News C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
RA O'neill K., Osman S., Parker S., Perrin D., Phunkhang P., Piquani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Ketta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thoulutsang Y., Topham K.,
RA Towey S., Tsamla T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,

RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Ustilago maydis";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.

DR EMBL; AACP01000173; EAK86150.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 247 AA; 27148 MW; 015A754D369D7EF2 CRC64;

Query Match 83.6%; Score 56; DB 2; Length 247;
Best Local Similarity 80.0%; Pred. No. 86;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYP PPPPIP 10
||| |||||:
Db 141 PPPPP PPPPVP 150

RESULT 37

Q5FVW5 XENTR PRELIMINARY; PRT; 353 AA.

AC Q5FVW5;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE LOC548379 protein (Fragment).
GN Name=LOC548379;
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Silurana.
OX NCBI_TaxID=8364;
RN [1]

RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Whole body;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Touchman J.W., Green E.D., Dickson M.C.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]

RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Whole body;
RA Klein S., Gerhard D.S.;
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC089740; AAH89740.1; -; mRNA.
FT NON TER 1 1
SQ SEQUENCE 353 AA; 39442 MW; E731FF294870F9EC CRC64;

Query Match 83.6%; Score 56; DB 2; Length 353;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYP PPPPIP 10
||| |||||:
Db 174 PPPPP PPPPVP 183

```
RESULT 38
EVL_RAT
ID EVL_RAT STANDARD; PRT; 393 AA.
AC O08719;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Ena/vasodilator stimulated phosphoprotein-like protein (Ena/VASP-like
DE protein).
GN Name=Evl; Synonyms=Rnb6;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Murioidea; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE, TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
RC STRAIN=Wistar; TISSUE=Brain;
RX MEDLINE=97415794; PubMed=9268706; DOI=10.1006/bbrc.1997.7113;
RA Ohta S., Mineta T., Kimoto M., Tabuchi K.;
RT "Differential display cloning of a novel rat cDNA (RNB6) that shows
RT high expression in the neonatal brain revealed a member of Ena/VASP
RT family.";
RL Biochem. Biophys. Res. Commun. 237:307-312(1997).
CC -!- FUNCTION: Enhances actin nucleation and polymerization (By
CC similarity).
CC -!- SUBUNIT: Binds to the SH3 domains of ABL1, LYN and SRC. Also binds
CC to profilin and the WW domain of APBB1/FE65. Binds to SEVA6A (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Focal adhesions (By similarity).
CC -!- TISSUE SPECIFICITY: Expression detected in brain, spleen, thymus
CC and testis.
CC -!- DEVELOPMENTAL STAGE: In the brain, expression gradually increases
CC during embryonic development, reaches a maximum at postnatal day 1
CC and decreases thereafter.
CC -!- PTM: Phosphorylated by PKA; phosphorylation abolishes binding to
CC SH3 domains of ABL and SRC (By similarity).
CC -!- SIMILARITY: Contains 1 WH1 domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; U70211; AAC53322.1; -; mRNA.
CC FIR; JC5614; JC5614.
CC HSSP; P70429; 1QC6.
CC SMR; O08719; 1-113.
CC Ensembl; ENSRNOG00000014476; Rattus norvegicus.
CC RGD; 621150; Evl.
CC GO; GO:0005737; C:cytoplasm; ISS.
CC GO; GO:0005925; C:focal adhesion; ISS.
CC GO; GO:0030027; C:lamellipodium; ISS.
CC GO; GO:0005522; F:profilin binding; ISS.
CC GO; GO:0017124; F:SH3 domain binding; ISS.
CC GO; GO:0030048; P:actin filament-based movement; ISS.
CC GO; GO:0045010; P:actin nucleation; ISS.
CC GO; GO:0008154; P:actin polymerization and/or depolymerization; ISS.
CC GO; GO:0007411; P:axon guidance; NAS.
CC GO; GO:0051016; P:barbed-end actin filament capping; ISS.
CC GO; GO:0006928; P:cell motility; NAS.
CC GO; GO:0007166; P:cell surface receptor linked signal transdu. . .; ISS.
CC GO; GO:0007417; P:central nervous system development; TAS.
CC GO; GO:0009887; P:organogenesis; ISS.
CC GO; GO:0030168; P:platelet activation; ISS.
CC InterPro; IPR000697; EVH1.
CC InterPro; IPR011993; PH_type.
CC Pfam; PF00568; WH1; 1.
CC SMART; SM00461; WH1; 1.
CC PROSITE; PS50229; WH1; 1.
```

```
KW Phosphorylation; SH3-binding. WH1.
FT DOMAIN 1 112 Pro-rich.
FT COMPIAS 160 204 Phosphoserine (By similarity).
FT MOD_RES 327 329 Phosphoserine (By similarity).
FT MOD_RES 329 329 Phosphoserine (By similarity).
SQ SEQUENCE 393 AA; 42095 MW; 6371D91362925D4E CRC64;

Query Match 83.6%; Score 56; DB 1; Length 393;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPPYP PPPIP 10
Db 180 PPPPP PPPVP 189
||| ||||:|
||| ||||:|

RESULT 39
Q6PB99 MOUSE
ID Q6PB99 MOUSE PRELIMINARY; PRT; 393 AA.
AC Q6PB99;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Ena-vasodilator stimulated phosphoprotein.
GN Name=Evl;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RG NIH MGC Project;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC059810; AAH59810.1; -; mRNA.
DR SMR; Q6PB99; 1-113.
DR MGI; MGI:1194884; Evl.
DR GO; GO:0005515; F:protein binding; IPI.
DR InterPro; IPR000697; EVH1.
DR InterPro; IPR001960; WH1.
DR Pfam; PF00568; WH1; 1.
DR SMART; SM00461; WH1; 1.
SQ SEQUENCE 393 AA; 42093 MW; A76FD82A2B415D8D CRC64;

Query Match 83.6%; Score 56; DB 2; Length 393;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

QY 1 PPPYPPPPPIP 10
Db 180 PPPPPPPPPVP 189

RESULT 40
Q501K3 MOUSE
ID Q501K3_MOUSE PRELIMINARY; PRT; 405 AA.
AC Q501K3;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Evl protein.
GN Name=Evl;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Kleusner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RG NIH MGC Project;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BC096017; AAH96017.1; -; mRNA.
DR SMR; Q501K3; 1-113.
DR MGI; MGI:1194884; Evl.
DR GO; GO:0005515; F:protein binding; IPI.
DR InterPro; IPR000697; EVH1.
DR InterPro; IPR011988; MHCII_invariant.
DR InterPro; IPR011993; PH_type.
DR InterPro; IPR001960; WH1.
DR Pfam; PF00568; WH1; 1.
DR SMART; SM00461; WH1; 1.
SQ SEQUENCE 405 AA; 43420 MW; 8D76F377139534EC CRC64;

Query Match 83.6%; Score 56; DB 2; Length 405;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
Db 180 PPPPPPPPPVP 189

RESULT 41
Q9SBM1 VOLCA
ID Q9SBM1_VOLCA PRELIMINARY; PRT; 409 AA.
AC Q9SBM1;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hydroxyproline-rich glycoprotein DZ-HRGP precursor.
GN Name=HRGP gene;
OS Volvox carteri f. nagariensis.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae;
OC Chlamydomonadales; Volvocaceae; Volvox.
OX NCBI_TaxID=3068;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HK10;
RX MEDLINE=20044763; PubMed=10574980; DOI=10.1074/jbc.274.49.35023;
RA Ender F., Hallmann A., Amon P., Sumper M.;
RT "Response to the sexual pheromone and wounding in the green alga
volvox: induction of an extracellular glycoprotein consisting almost
exclusively of hydroxyproline.";
RL J. Biol. Chem. 274:35023-35028(1999).
DR EMBL; AJ242540; CAB62280.1; -; mRNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005199; F:structural constituent of cell wall; IEA.
DR GO; GO:0005249; F:voltage-gated potassium channel activity; IEA.
DR GO; GO:0006813; P:potassium ion transport; IEA.
DR InterPro; IPR002951; Atrophin.
DR InterPro; IPR005404; KV3.3 channel.
DR PRINTS; PR01222; ATROPHIN.
DR PRINTS; PR01582; KV33CHANNEL.
DR PRINTS; PR01218; PSTLEXTENSIN.
KW Signal.
FT SIGNAL 1 18 Potential.
FT CHAIN 18 409 hydroxyproline-rich glycoprotein DZ-HRGP.
SQ SEQUENCE 409 AA; 41547 MW; CD0749C6AF02BD74 CRC64;

Query Match 83.6%; Score 56; DB 2; Length 409;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPPIP 10
Db 87 PPPPPPPPPVP 96

RESULT 42
EVL_MOUSE
ID EVL_MOUSE STANDARD; PRT; 414 AA.
AC P70429; Q9ERU8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Ena/vasodilator stimulated phosphoprotein-like protein (Ena/VASP-like
protein).
DE Name=Evl;
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
RC TISSUE=Brain;
RX MEDLINE=97015079; PubMed=8861907; DOI=10.1016/S0092-8674(00)81341-0;
RA Gertler F.B., Niebuhr K., Reinhard M., Wehland J., Soriano P.;
RT "Mena, a relative of VASP and Drosophila Enabled, is implicated in the
control of microfilament dynamics.";
RL Cell 87:227-239(1996).
RN [2]
RP NUCLEOTIDE SEQUENCE (ISOFORM 2), FUNCTION, SUBUNIT, SUBCELLULAR
LOCATION, AND PHOSPHORYLATION BY PKA.
RC STRAIN=C57BL/6J;
RX PubMed=10945997; DOI=10.1074/jbc.M006274200;
RA Lambrechts A., Kwiatkowski A.V., Lanier L.M., Bear J.E.,
RA Vandekerckhove J., Ampe C., Gertler F.B.;

RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 163-364.
RC TISSUE=Uterus;
RG The German CDNA consortium;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [6]
RP INTERACTION WITH SEMA6A.
RX PubMed=10993894; DOI=10.1074/jbc.M006316200;
RA Klostermann A., Lutz B., Gertler F., Behl C.;
RT "The orthologous human and murine semaphorin 6A-1 proteins (SEMA6A-1/Sema6A-1) bind to the enabled/vasodilator-stimulated phosphoprotein-like protein (EVL) via a novel carboxyl-terminal zyxin-like domain."; J. Biol. Chem. 275:39647-39653 (2000).
RL [7]
RN PHOSPHORYLATION SITES SER-329 AND SER-331.
RX PubMed=15302935; DOI=10.1073/pnas.0404720101;
RA Beausoleil S.A., Jedrychowski M., Schwartz D., Elias J.E., Villen J., Li J., Cohn M.A., Cantley L.C., Gygi S.P.;
RT "Large-scale characterization of HeLa cell nuclear phosphoproteins."; Proc. Natl. Acad. Sci. U.S.A. 101:12130-12135 (2004).
CC -!- FUNCTION: Enhances actin nucleation and polymerization (By similarity).
CC -!- SUBUNIT: Binds to the SH3 domains of ABL1, LYN and SRC. Also binds to profilin and the WW domain of APBB1/FE65 (By similarity). Binds to SEMA6A.
CC -!- INTERACTION:
CC P08631:HCK; NbExp=1; IntAct=EBI-346653, EBI-346340;
CC -!- SUBCELLULAR LOCATION: Focal adhesions (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=2;
CC IsoId=Q9UI08-1; Sequence=Displayed;
CC Name=1;
CC IsoId=Q9UI08-2; Sequence=VSP_004044;
CC -!- PTM: Phosphorylated by PKA; phosphorylation abolishes binding to SH3 domains of ABL and SRC (By similarity).
CC -!- SIMILARITY: Contains 1 WH1 domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.
CC -----
DR EMBL; AF052504; AAF21709.1; -; mRNA.
DR EMBL; AF112209; AAF17197.1; -; mRNA.
DR EMBL; BC023997; AAH23997.1; -; mRNA.
DR EMBL; BC032358; AAH32358.1; -; mRNA.
DR EMBL; AF131766; AAD20040.1; -; mRNA.
DR EMBL; AL133642; CAB63763.2; -; mRNA.
DR HSSP; P70429; 1QC6.
DR SMR; Q9UI08; 1-113.
DR IntAct; Q9UI08; -.
DR Ensembl; ENSG00000196405; Homo sapiens.
DR HGNC; HGNC:20234; EVL.
DR H-InvDB; HIX0011964; -.
DR GO; GO:0005737; C:cytoplasm; ISS.
DR GO; GO:0005925; C:focal adhesion; ISS.
DR GO; GO:0030027; C:lamellipodium; ISS.
DR GO; GO:0005522; F:profilin binding; ISS.
DR GO; GO:0017124; F:SH3 domain binding; ISS.
DR GO; GO:0007015; P:actin filament organization; TAS.
DR GO; GO:0030048; P:actin filament-based movement; ISS.
DR GO; GO:0045010; P:actin nucleation; ISS.
DR GO; GO:0008154; P:actin polymerization and/or depolymerization; ISS.
DR GO; GO:0007411; P:axon guidance; ISS.
DR GO; GO:0051016; P:barbed-end actin filament capping; ISS.
DR GO; GO:0007166; P:cell surface receptor linked signal transdu. . .; NAS.
DR GO; GO:0007399; P:neurogenesis; NAS.
DR GO; GO:0030168; P:platelet activation; ISS.
DR InterPro; IPR000697; EVH1.
DR InterPro; IPR011993; PH_type.

DR Pfam; PF00568; WH1; 1.
DR SMART; SM00461; WH1; 1.
DR PROSITE; PS50229; WH1; 1.
KW Alternative splicing; Phosphorylation; SH3-binding.
FT DOMAIN 1 112 WH1.
FT COMPBIAS 162 206 Pro-rich.
FT MOD_RES 329 329 Phosphoserine.
FT MOD_RES 331 331 Phosphoserine.
FT VARSPLIC 1 1 M -> MAT (in isoform 1).
FT CONFLICT 201 201 /FTId=VSP_004044.
FT CONFLICT 364 364 p -> S (in Ref. 2).
FT CONFLICT 364 364 M -> Y (in Ref. 5).
SQ SEQUENCE 416 AA; 44620 MW; AD5B67458755D659 CRC64;

Query Match 83.6%; Score 56; DB 1; Length 416;
Best Local Similarity 80.0%; Pred. No. 1.5e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYP PPPPIP 10
||| ||||:|
Db 182 PPPPP PPPVP 191

RESULT 44
Q7Z522 HUMAN
ID Q7Z522_HUMAN PRELIMINARY; PRT; 418 AA.
AC Q7Z522;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE B6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Wan Y.Z., Yu L., Yue P., Tu Q., Fu S.N., Zhao S.Y.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Zhang M., Yu L., Zhou Y., Hu P.R., Xin Y.R., Zhao S.Y.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF087843; AAP97156.1; -; mRNA.
DR HSSP; P50552; 1EGX.
DR SMR; Q7Z522; 4-115.
DR Ensembl; ENSG00000196405; Homo sapiens.
DR InterPro; IPR000697; EVH1.
DR InterPro; IPR011993; PH type.
DR InterPro; IPR001960; WH1.
DR Pfam; PF00568; WH1; 1.
DR SMART; SM00461; WH1; 1.
SQ SEQUENCE 418 AA; 44819 MW; 02D02741E5EE7219 CRC64;

Query Match 83.6%; Score 56; DB 2; Length 418;
Best Local Similarity 80.0%; Pred. No. 1.5e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYP PPPPIP 10
||| ||||:|
Db 184 PPPPP PPPVP 193

RESULT 45
QSR896 PONPY
ID QSR896_PONPY PRELIMINARY; PRT; 422 AA.
AC QSR896;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein DKFZp468J0419.
GN Name=DKFZp468J0419;

Best Local Similarity 80.0%; Pred. No. 1.7e+02; Mismatches 1; Indels 0; Gaps 0;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 PPPYPPPPIP 10
Db 251 PPPSPPPPLP 260

RESULT 49
Q73XK9 MYCPA
ID Q73XK9_MYCPA PRELIMINARY; PRT; 489 AA.
AC Q73XK9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=MAP2300C;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB017235; AAS04617.1; -; Genomic_DNA.
DR InterPro; IPR004255; UPF0089.
DR Pfam; PF03007; UPF0089; 1.
KW Complete proteome.
SQ SEQUENCE 489 AA; 52083 MW; 40A98903DB6D6890 CRC64;

Query Match 83.6%; Score 56; DB 2; Length 489;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPPYPPPPIP 10
Db 167 PPPKPPPPPLP 176

RESULT 50
Q4R7B7 MACFA
ID Q4R7B7_MACFA PRELIMINARY; PRT; 557 AA.
AC Q4R7B7;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Testis cDNA, clone: QtsA-15671, similar to human hypothetical protein LOC195977 (LOC195977).
DE Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Cercopithecidae; Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA International consortium for macaque cDNA sequencing, analysis; "DNA sequences of macaque genes expressed in brain or testis and its evolutionary implications.";
RT Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Osada N., Hirata M., Tanuma R., Kusuda J., Hida M., Suzuki Y., Sugano S., Gojobori T., Shen J.C.-K., Wu C.I., Hashimoto K.;
RT "Substitution rate and structural divergence of 5'UTR evolution: Comparative analysis between human and cynomolgus monkey cDNAs.";
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB168903; BAE01005.1; -; mRNA.
KW Hypothetical protein.
SQ SEQUENCE 557 AA; 62019 MW; B524252DEBE18FAC CRC64;

Query Match 83.6%; Score 56; DB 2; Length 557;
Best Local Similarity 80.0%; Pred. No. 2e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 PPPYPPPPIP 10
Db 400 PPPSPPPPLP 409

Search completed: April 6, 2006, 09:38:59
Job time : 172.895 secs